On the Science of Grasping
Modelling Grasp Affordances in Robotics from Human Analysis

Cotugno, Giuseppe

Awarding institution:
King's College London

KING'S College LONDON

On the Science of Grasping: Modelling Grasp Affordances in Robotics from Human Analysis

by Giuseppe Cotugno

Supervisors:

Dr. Thrishantha Nanayakkara

Prof. Kaspar Althoefer

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Robotics

to

King's College London
School of Natural and Mathematical Sciences
Department of Informatics
Centre for Robotics Research

June 2018

I herewith declare that I have produced this PhD thesis without the prohibited assistance of third parties and without making use of aids other than those specified; notions taken over directly or indirectly from other sources have been identified as such. This PhD thesis has not previously been presented in identical or similar form to any other British or foreign examination board.

London, June 2018

This thesis is dedicated to my dear daughter Anastasia and my beloved wife Liza. Omnia vincit amor.

Abstract

This thesis investigates the human motion patterns of grasping and approaching, as well as applies the functional features to a robotic grasp affordances controller.

Human grasping has always been seen as a target milestone in robotics; however, it is still difficult for robots to perform advanced grasping tasks, such as required in an assembly line. For robots to operate in an unstructured environment, it is necessary to focus the attention on performing a specific action with any available object by grasping it in the most proficient way. The neuroscientific application of the theory of affordances explains how this process happens for humans.

This thesis analyses human behaviours of grasping and approaching to grasp when performing an action. This study is needed to understand the common behavioural and morphological factors that influence the motion planning and selection of grasp postures and grasp affordances. The main features of motion are statistically analysed and adapted to robotic application. A human inspired approaching and grasping controller is proposed, and the limitations and benefits of state of the art robotic hands are derived based on human data.

The approach of this thesis differs from the traditional method of studying grasp affordances. Different elements of the problem, such as perception, learning, motor control and intention, are modularly separated in components rather than monolithically aggregated in a single entity. A biologically motivated modular grasp affordance system can adapt only selected features of human grasping to robotics, leaving out biological elements not needed in robotic applications. The modularity of the system separates cognitive motor decisions from the robot's own embodiment and the geometric properties of the objects, granting more independence from the specific domain of application.

Acknowledgements

This PhD has been a long and troubled journey, from the beginning of its conception to the completion of the studies. As such, this work has been made possible with the professional and personal help and contribution of many people.

First of all I wish to thank Dr. Thrishantha Nanayakkara, for his significant help and contribution in defining the scientific ideas behind this thesis and maturing my own scientific principles and methodology presented in this manuscript. I express my gratitude to Prof. Kaspar Althoefer, for his advices and financial support, and Dr. Elizabeth Sklar for her assistance in the last stage of this work. I thank Dr. Jelizaveta Konstantinova, for her fundamental assistance, scientific support and guidance in the crucial moments of this work.

I also acknowledge the contribution of the following European research grants which funded these studies: Framework Program 7 DARWIN Project (Grant No. 270138), which funded the initial part of this work, Horizon 2020 4x3 Project (Grant No. 637095) and Framework Program 7 SQUIRREL Project (Grant No. 610532).

I am also very grateful to Prof. John Gerald Taylor: his scientific ideas influenced profoundly the methodology adopted in this thesis and engaged me in the study of human neuroscience. I thank Prof. Leonard White, for freely disseminating his knowledge of brain studies, and Prof. Roberta Sala for her help for the literature search for the conclusions.

I acknowledge the moral and professional support of my friends and colleagues, in the DARWIN project and both at King's College London and at Buhler UK: Dr. Kris De Meyer, Dr. Dimitri Ognibene, Dr. Vishwanathan Mohan, Dr. Manolis Lourakis, Dr. Xenophon Zabulis, Prof. Jiri Matas, Dr. Tomas Hodan, Sharat Akkaladevi, Dr. Anuradha Ranasinghe, Dr. Georgia Kyriakidou, Dr. Min Li, Dr. Angela Faragasso, Dr. Joao Bimbo, Dr. Helge Wudermann, Dr. Agostino Stilli, Prof. Harris Makatsoris, Rob Davis - who made me a better engineer, Guangqing You, Foreed Choudhury, Tony Hug, Dr. Jean-Francois Deprez, Jim Cameron, Steve Newton, Jamie Mitchinson, Hemant Kapoor, Ram Mulage, John Brown, James Goldsmith and Duncan Shepard. I also acknowledge the support of my friends in Italy: Luca Tosti, Fabio, Veronica and Riccardo Capasso, Alessandro Gallucci, Silvio and Manuela Pierro.

At the very last, I recognise the vital help of the most important people, that made all this possible: my dearest wife Liza, for her patience, understanding and advice, and my beloved daughter Anastasia. My deepest gratitude goes to my dad and mum, Vittorio and Alba, my father and mother in law, Sergei and Natalija, my sisters Adelaide and Alessandra with her husband, Asad, and son, Christopher, and my nephew Ricard. I am grateful to have all of you on my side as you made all this possible with you love, encouragements and unconditioned assistance.

The original question, "Can machines think?" I believe to be too meaningless to deserve discussion.

Nevertheless I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted.

Alan M. Turing, "Computing machinery and intelligence". *Mind*, 59, 433-460, 1950.

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

Introduction

1.1 Motivation of the Thesis

The ability of humans to manipulate objects and craft tools gives us the possibility to shape and significantly alter the environment. Those skills were developed during the course of human evolution, and reached its apogee with the crafting of the Oldowan stone tools by the *Homo Habilis* [1]. As primitive humans improved their abilities in crafting tools and altering the environment, their cognitive skills were improving [2] and more advanced tools and objects were created to help performing daily tasks and duties. As a result, during the course of history, all the man-made objects were produced taking into account the morphology of human hands.

Recent advancements in robotics research have made realistic the possibility of employing robotic assistants in household and industrial environments. Commercial applications of robots performing household duties, such as lawn mowing[1] or vacuum cleaning[2] are already distributed to the wider public, and industrial set-ups of robots performing

[1]Bosch GmbH, Germany
[2]iRobot Corporation, USA

flexible industrial tasks are increasingly being deployed[3]. As robotics applications are gaining momentum, it is fundamental for a robotic system to be able to interact with man-made objects and tools. However, such ability to interact with and to manipulate the environment is still far from being achieved.

Although the problem of object manipulation was formulated for the first time in the 1980s [3], it is still considered a research problem to achieve a level of dexterity that is suitable to employ robots in unstructured environments[4] or when a fast response to unexpected events in industrial set-ups[5] is required. Moreover, some tasks are difficult to automate due to the high variability of the features of the objects involved. Current grasping solutions allow the use of a limited number of objects and, as a result, the existing manipulation controllers tend to be tailored to the specific problem [4]. For this reason, a more comprehensive approach to manipulation, based on object affordances is becoming increasingly popular. The concept, first formulated in psychology by Gibson in 1977 [5], states that an object affords different uses for itself based on its geometrical features. For instance, a spoon affords to be used to collect food, to mix a liquid or to hammer on a point. This idea was developed further in neuroscience by introducing the concept of *grasp affordances*. They describe the different strategies to grasp the same object and play a role in neural control of volitional motion [6], as well as the selection of appropriate grasping strategies [7]. Object and grasp affordances have the potential to improve the manipulation skills of robots. Grasp affordances could exploit the functionalities of existing hardware for approaching and grasping an object, leaving tailored solutions as an optional choice. Additionally, object affordances can be used to derive the best manipulation plan that executes an elaborated task, such as an assembly, even if tools are missing or locations of objects are not specified.

[3]Loop Technology LTD, UK; Profactor GmbH, Austria

[4]European Community's Seventh Framework Programme (FP7/2014-2018) project SQUIRREL (Grant No: FP7-610532)

[5]European Community's Seventh Framework Programme (FP7/2007-2013) project DARWIN (Grant No: FP7-270138)

The idea of a robot being able to cook or to assemble a furniture is very appealing, and the available computational power allows implementing sophisticated hand and arm planning algorithms. Despite this, the manipulation capabilities of the existing robotic hardware are too limited to benefit from the advanced cognitive capabilities of object affordances [8,9]. Indeed, an object or tool has to be grasped solidly and appropriately before being correctly used and, until such skill has not been perfected, it is difficult to fully exploit the potentialities of object affordances. Therefore, it is first required to improve the existing grasping and manipulation control using the possibilities of available robotic hands and arms.

Traditionally, the different definitions of grasp affordances in robotics combine many aspects of the problem - from perception to grasping. They are defined in a single inseparable formulation that does not scale well for different objects or tasks [10]. The study of a modular formulation of grasp affordances in robotics, that separates each part of the problem, such as perception from learning, and is inspired by human motor control, is a required initial step towards robotic manipulation control. Such approach is independent from the features of the objects used in a specific domain of application. The study of human motor control is required to understand the principles that made human manipulation so versatile. The exploration of those principles is not targeted at the creation of a mere emulation of human manipulation capabilities. It is rather aimed to define the main features of human manipulation in order to identify the principles that should be used in robotics. Following this approach, the aim of this thesis is to advance the state of the art of robotic manipulation by developing the concept of *grasp affordances* in robotics in a modular formulation that separates different aspects of the problem. Each component is derived from human studies and is implemented on a robotic platform as a proof of concept.

1.2 Aim and Objectives of the Thesis

The aim of this thesis is *to characterise the grasping and approaching parts of human grasp affordances and to translate and adapt those principles to robotic manipulation.* The objectives of the thesis are:

1. To contribute to the definition of a modular design of a grasp affordance controller, where different parts, such as perception, grasping and reaching, are separated in different components;

2. To characterise patterns and features of human grasping and approaching motions that are independent from the manipulated object;

3. To translate and adapt the detected and existing biological principles of manipulation to robotic grasp affordance control, taking into account the differences between human and robot manipulators in kinematics, dexterity and application;

4. To propose a formulation of the problem of grasp affordances that depends on the goal of the action undertaken, rather than the specific object properties.

1.3 Contributions

The contributions of the thesis can be summarised as follows:

1. Grasp affordances can be decomposed in two parallel actions: reaching to the object while shaping the grip (approaching) and adjusting the hand posture when completing the grasp. This thesis characterises the approaching part of a grasp affordance by relating hand and object in a spherical coordinate system. The analysis of human studies showed that: 1) the control of the hand-to-object distance

follows an object-independent pattern of defined and is common across subjects, 2) the control of the hand-to-object orientation is a planned action, 3) the independent control of the hand rotation follows a reactive motion. Models for each examined quantity and phase of the motion are derived and validated (Chapter 3).

2. The high dexterity of the thumb is one of the factors that grants advanced manipulation capabilities to humans, however, its role in grasping is not quantified. The characteristics of the displacement of the fingers of the human hand are statistically evaluated in terms of variability of the motion. It is demonstrated that the motion of the thumb leads the displacement of the other fingers and determines the specific grasp posture being used. The kinematic mismatch between robotic and human thumbs is quantified and the implications of such limitation on grasp affordances are discussed. Suggestions for an improved kinematic design of robotic hands are also outlined (Chapter 4).

3. The proposed architecture of modular grasp affordance control requires to be validated as a proof of concept. The models of the approaching to grasp strategies, outlined in this thesis, are implemented on a simulated robotic platform as modular components and their performance and effectiveness are evaluated. The feasibility and grasping performance of the resulting system is assessed against action-dependent set criteria on the example of a hammering action (Chapter 5).

4. A suitable grasping controller for multifingered hands is required for a complete grasp affordance system in robotics. A software controller is proposed, which takes advantage of the principles defined by the role of the thumb in grasping. The controller learns an initial grasp posture from kinaesthetic demonstrations and is able to manipulate everyday objects of different shapes and unrelated to the original learned set (Chapter 6).

1.4 Contributed Papers

The findings of this thesis were published in the following peer reviewed international journal and conference papers:

Journal Papers

1. G. Cotugno, J. Konstantinova, K. Althoefer, T. Nanayakkara, "Modelling the Structure of Object-Independent Human Affordances of Approaching to Grasp for Robotic Hands", *PLOS One*, under review

2. G. Cotugno, K. Althoefer, T. Nanayakkara, "The Role of the Thumb: Study of Finger Motion in Grasping and Reachability Space in Human and Robotic Hands", *IEEE Transactions in Systems, Man and Cybernetics: Part A*, Vol. 47, Issue 7, pp 1061 - 1070, July 2017

Conference Papers

1. G. Cotugno, J. Konstantinova, A. Stilli, Y. Noh, K. Althoefer, "Object Classification Using Hybrid Fiber Optical Force/Proximity Sensor", in *Proceedings on IEEE Sensors 2017*, 30 October - 1 November, 2017

2. G. Cotugno, V. Mohan, K. Althoefer, T. Nanayakkara, "Simplifying Grasping Complexity through Generalization of Kinaesthetically Learned Synergies", in *IEEE International Conference on Robotics and Automation*, pp. 5345 - 5351, May 31 - June 7, 2014

3. G. Cotugno, J. Konstantinova, K. Althoefer, T. Nanayakkara, "On the dexterity of robotic manipulation: are robotic hands ill designed?", in the *6th International Conference on Cognitive Science*, 23 - 27 June 2014

4. G. Cotugno, A. Ibrahim, K. Althoefer, T. Nanayakkara, "Human affordances of stacking: Best placement or best outlook?", in the *6th International Conference on Cognitive Science*, 23 - 27 June 2014

5. G. Cotugno, L. D'Alfonso, W. Lucia, P. Muraca, P. Pugliese, "Extended and Unscented Kalman Filters for mobile robot localization and environment reconstruction", in *21st IEEE Mediterranean Conference on Control and Automation*, 25 - 28 June 2013

Co-authored Papers

1. V. K. Nanayakkara, *G. Cotugno*, N. Vitzilaios, D. Venetsanos, T. Nanayakkara, and M. Necip Sahinkaya, "The Role of Morphology of the Thumb in Anthropomorphic Grasping: A Review", in Frontiers in Mechanical Engineering, section Mechatronics, June 2017

2. J. Konstantinova, *G. Cotugno*, P. Dasgupta, K. Althoefer and T. Nanayakkara, "Palpation force modulation strategies to identify hard regions in soft tissue organs", in PLoS ONE, February 2017

3. J. Konstantinova, *G. Cotugno*, P. Dasgupta, K. Althoefer and T. Nanayakkara, "Autonomous Robotic Palpation of Soft Tissue using the Modulation of Applied Force", in 6th IEEE International Conference on Biomedical Robotics and Biomechatronics, 26 - 29 June 2016

4. W. Torgerud, D. Mussack, T. Lee, G. Maffei, *G. Cotugno* and P. Schrater, "Arousal decreases conservativeness in a random dot motion decision making task", in Neuroscience - Society for Neuroscience, 17 - 21 October 2015

5. W. Torgerud, D. Mussack, T. Lee, G. Maffei, *G. Cotugno* and P. Schrater, "Affective decision making: effects of arousal on a random dot motion task", in

Computational and Systems Neuroscience, 5 - 8 March 2015

1.5 Thesis Structure

The contributions of this work are divided in the following chapters:

Chapter 2 provides context to this work explaining the methodology and the challenges addressed in robotics and how this work advances the state of the art of grasp affordances in robotics. For this purpose, an introduction to the neural control mechanism of voluntary movements, such as grasping, and to the anatomical and morphological features of the human hand is given.

Chapter 3 introduces object independent models of grasp affordances of hammering derived from human demonstrations. The approaching motion is characterised in terms of distance and orientation displacements between the object and the hand.

Chapter 4 characterise the fundamental role of the thumb when robustly gripping different everyday objects. It also quantifies the kinematic mismatch between robotic and human thumb reachability spaces.

Chapter 5 presents a modular architecture of a grasp affordance system, and formulates the grasp affordance problem in terms of relationships between the robot and parts of the object. The algorithm approaching trajectory generator, able to employ different human models of approaching, is described. The feasibility of the system is validated and compared with a state of the art planner in a simulation study of a humanoid robot replicating an hammering approaching to grasp motion.

In **Chapter 6** the principal role of the thumb in grasping is applied to a state-of-the-art robotic multifingered hand to grasp different objects from a kinaesthetically learned

initial posture. The biological principles described in the previous chapters are applied to the control of real robots for manipulation of different objects.

Chapter 7 summarises this work and concludes it by describing future directions, limitations and possibilities for improvement.

1.6 Thesis Map

The structure of the thesis is schematically shown on Fig. 1.1.

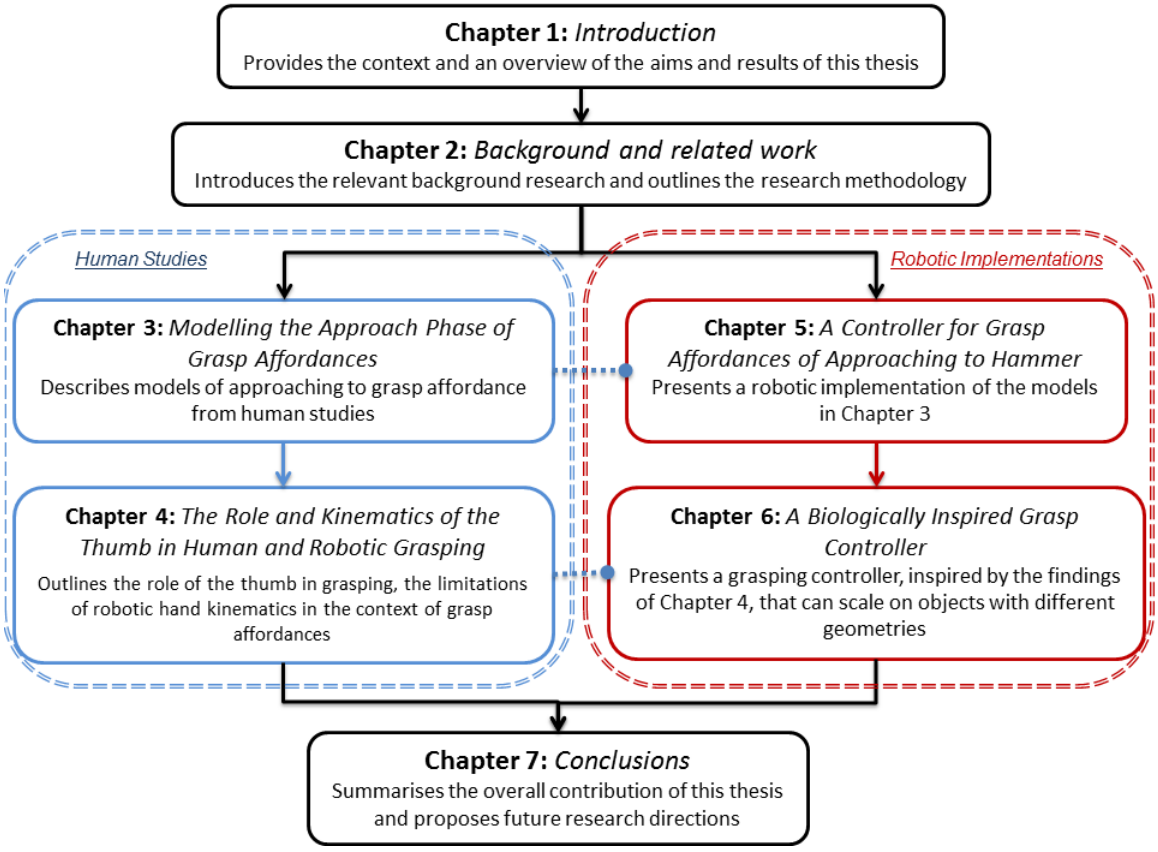


Figure 1.1: Schematic structure of the thesis

Chapter 2

Background and Related Work

2.1 Introduction

Most of daily household or industrial tasks require some level of interaction with the environment. Excluding some operations such as the delivery of parcels or surveillance, for most of the tasks an operator manipulates the environment directly, such as when harvesting crops, using tools in an assembly line, or while cleaning a house. Since the introduction of the steam engine in the late XVIII century, automation had an increasing role in the execution of formerly manual tasks, increasing the throughputs and reducing the effort and time devoted by a human operator as technology improved [11]. Nowadays many industries, such as agriculture or car manufacturing, heavily rely on machines and automation. However the limited flexibility and adaptation of the present mass production processes, generates costly inefficiencies. Unsold items are wasted or disposed, creating a loss for the manufacturer and, potentially, for the environment since most of items cannot be produced on demand. One of the reasons for the lack of production flexibility is that many tasks are still performed mostly by humans using specialised tools, such as in landmine removals [12] or in assemblies of customer-tailored

products (*just-in-time* assemblies [13]). Performing tasks mostly manually has many disadvantages, such as low production yields, the risk of injury or death, and the poor job quality which alienates and demotivates the workers in the long run [14]. The lack of flexibility of modern industrial set-ups requires an increased dependence on inexpensive manual labour to reduce the production costs and adapt the product to the customer's needs. It is, therefore, highly desirable to decrease the use of manual work by introducing robots able to collaborate with human labour. This can be a mean to produce customised, on demand items at a low cost reducing the environmental waste created when disposing unsold objects. The use of robotic assistants can improve the job quality and safety of the work environment, as humans do not require performing repetitive, heavy or dangerous tasks. For this to be possible, it is required to reproduce the advanced manipulation capabilities of humans for robotic manipulation.

Humans gained their advanced manipulation capabilities during the course of evolution, as they were able to develop and control the opposable thumb among the other fingers, between 2.4 and 1.4 million years ago [15]. The thumb gives a great contribution to manipulation, and its loss can severely impair the dexterity of the hand [16]. Indeed, the only animals that exhibit manipulation skills are human and non-human primates [17], which all have a hand with an opposable thumb. Different morphologies of humanoid hands still can grant some level of dexterity, as it is the case of non-human primates such as capuchins [18] and chimpanzees [19]. Although those primates can manipulate the environment, humans still show a finer control of the fingers, as the set of postures that they are able to adopt is more advanced [18] when compared with primates [20]. It can be, therefore, concluded that the manipulation skills of humans shall be taken as a reference, and their salient features shall be better understood to guide future research.

Initial research on robotic object manipulation was performed for the first time in the early 1960s in the field of prostethics [21]. The 1980s seen the first fundamental theoret-

ical works [22], definitions [23], and robotic prototypes were built [3]. The importance of the problem was increasing, as it is highly desirable for a robot to interact with human-made household or industrial objects without the need of additional customisations. A robot is the result of an engineered design, not bound by constraints of biological evolution, and it should be altered to better fit the target application. As a result, a wide range of end effectors [24] was developed to address a specific manipulation problem, such as screwing, soldering or picking and placing. The use of ad-hoc end effectors improved the mass production process, but does not solve the challenge of flexible small-scale assemblies, where a large number of objects and tools should be exchanged and used. This is largely an unsolved problem, as it is demonstrated by the numerous research and academic competitions on the subject, as well as the lack of commercial general-purpose robotic manipulation systems.

To overcome those limitations, there is a recent growing interest in studying human affordances of objects for manipulation [25]. The term *affordance* has two different interpretations, one psychological and the other neuroscientific, which have been already introduced in Chapter 1 as object and grasp affordance respectively. In this dissertation, the definition of affordance is inspired by the neuroscientific interpretation. In this work, an affordance is defined as *a possible way to approach and to grasp an object in order to perform a predefined action*. Understanding and adapting grasp affordances for robotics allows using existing hardware for generic domains, where multiple objects can be used to perform the same action. For example, this is the case when a furniture is assembled using different makes of the same tool. In this way, it can be possible to focus the attention on the actual plan of actions required to fulfil a goal, rather than on specific properties of the object being handled.

As humans have the most advanced manipulation skills available in nature, this chapter describes the morphology of human end effectors - the hands - and the known mech-

anisms of voluntary motor control. Voluntary, or volitional, motions are those motor actions which are intentionally performed, such as grasping or walking. By relating this to the current state of the art in robotics research, it is possible to understand which features are not well represented in robotics, and which ones are adequate or better than the human counterpart. From this discussion, a list of desirable requirements for robotic manipulation is outlined at the end of this chapter and addressed further in the technical chapters of this thesis. For this purpose, next sections describe both the biological and artificial systems in terms of a plant, as the musculoskeletal or electro-mechanic features of the hand, and the controller, as the neural or algorithmic mechanism actuating the plant.

In robotics scientific literature, it is rarely provided an overview of the mechanism of volitional motor control in humans, despite its relevance for grasping. Hence, the main contributions of the chapter is to provide such overview, aimed at a robotics audience, and to relate critically the mechanism of human motor control to robotics research.

2.2 Biological Background of Grasping

2.2.1 The Plant: The Human Hand

The human hand is a complex structure whose dexterity is unique in nature. Although its main well-known musculoskeletal features are invariable, there are still a lot of differences across subjects [27]. For instance, some sources of variability are associated with the size and length of the bones, others are related to the reachability space of each joint. An overview of the hand skeleton is shown in Figure 2.1. Specifically, some individuals are able to bend the distal intraphalangeal (DIP) joint of the thumb backwards (hitch-hiker's thumb) while others are not able to produce such motion. The skeleton

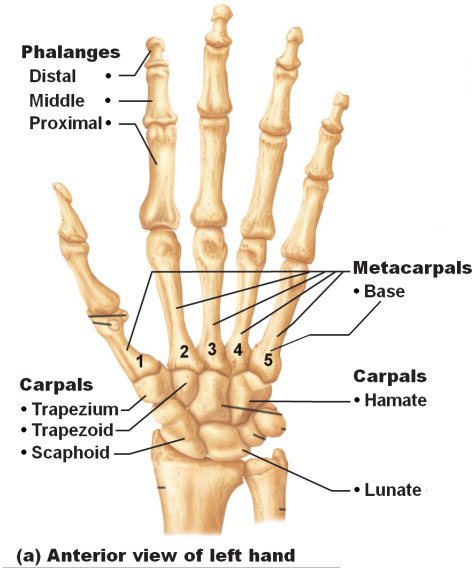


Figure 2.1: Description of the main bones of the human hand. Adapted from [26].

of the hand has 27 bones, which are divided in carpus, metacarpi and phalanxes [28]. The carpus is made of several small bones and is the lower base of the palm. It is connected and articulated with the radius, the forearm. The metacarpi link the carpus to the fingers and are part of the higher section of the palm up to the knuckles. The phalanxes are the bones of the five fingers of the hand, and can be grouped into proximal (close to the metacarpi), distal (at the tip of the fingers) and intra (between distal and proximal). The phalanxes of the thumb are only two: distal and proximal. The lightweight bone tissue grants elasticity and resistance to compression [29]. The skeleton is surrounded by fat, muscles and skin to provide enable friction, tactile, temperature and pain sensing capabilities. The muscles transduce electrical stimuli into a motion of the bone by pulling tendons. The muscles of the hand, and of the body, work in pair: as one contracts to initiate the motion, the other extends to relax. As such, muscles are grouped in agonist and antagonist couples, and most of the muscles of the fingers are concentrated on the forearm [28]. The thumb's muscles, instead, are positioned on the carpometacarpal complex of the hand itself. Across the muscles of the hand, the

muscles of the thumb are the strongest [30].

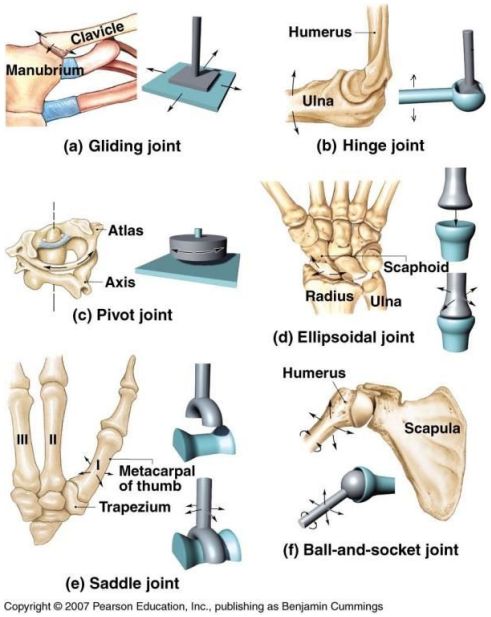


Figure 2.2: Description of the joint types found in the human skeleton. Hinge, ellipsoidal and saddle joints are found also in the hand skeleton. From [31].

When pulled by the muscles, the tendons position the bones and store motion energy, like a spring. Their elastic structure allows to store energy, like a spring, and resist stretches [32]. The muscle and tendon lengths balance each other, as longer muscles transduce larger forces while longer tendons promote more elastic motions, such as jumping or tapping [33]. Such differences also have a genetic explanation [34] and are a source of variability across subjects. The bones of the articulations are aligned by ligaments. Ligaments enable hand articulation and prevent the tendons to misplace a link incorrectly when translating the motion of the muscles [28]. The joints of the human skeleton can be divided into six categories based on the type of rotation they generate: gliding, ball and socket, hinge, pivot, ellipsoidal (or condyloid) and saddle. Figure 2.2 shows the possible rotations and some skeletal examples for each category. The pivot and the ball and socket joints are not found in the hand, while the ellipsoidal, the hinge and

the saddle joints articulate the metacarpals, the phalanxes and the thumb opposition respectively [28,35]. The opposition motion of the thumb is the most interesting, as it is not to rotating on a single axis. Indeed, this motion involves also the carpal bones, such as the trapezium, other than metacarpal and phalangeal bones, as the other fingers [36]. The precise mechanism of the thumb opposition is still matter of debate [37].

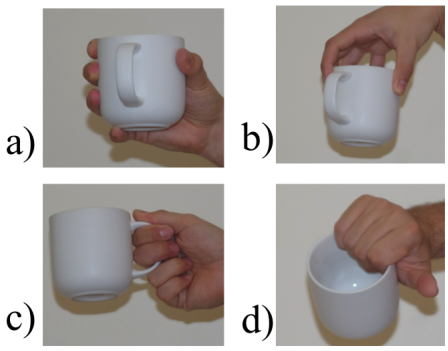


Figure 2.3: Grasping postures adopted by humans and primates: A) Power Grasp, B) Precision Grasp, C) Key Grasp, D) Grasping without using the thumb.

The oppositions of the thumb and little finger grant the ability to fold the palm and to further envelop an object, realising a category of grasp postures called *power grasps* [38], which is very popular among other primates too [18]. Alternatively, stable grasps can be obtained using the fingertips (*precision grasp* postures) [38] or using the thumb and the index as main supports (*key grasp* postures). The former posture category is more frequently seen in humans [18]. Finally, an object can be restrained using the sides or the bodies of the other fingers but the thumb, which is a strategy more used by other primates, such as the bonobo, the capuchin and new and old world monkeys in general [39]. The different grasp postures are pictured in Figure 2.3.

2.2.2 The Controller: Neuroscience of Volitional Motion

As mentioned, the musculoskeletal structure of the hand (the plant) is a complex system with great variabilities across subjects [27]. Nevertheless, humans are able to grasp and use everyday objects regardless of all the variabilities in the hand morphology. In order to control the hand and to make sure that objects are correctly grasped to perform the intended use, a distributed control system is employed as the controller by the human organism: the central nervous system (CNS) - the spinal cord and the brain - and its extensions.

The control system is distributed as its autonomy and intelligence is not retained only in the cerebral cortex, but it is scattered in different parts of the system. The cortex retains the responsibility of taking high-level decisions such as initiating a motion and prioritising a plan of actions. The neuroscience of motion is a very vast topic, which describes how reflexes, volitional motion and non-volitional motion (such as the heart-beat) of organs and muscles are controlled by the CNS. The CNS itself it is not the result of an engineered design, but it is a biological system that slowly evolved over the millennia. As such, its functionalities are not rigidly separated through well-defined interfaces, but they are overlapping and interleaved across the different parts of the CNS. For this reason, in this thesis, the description of the neural elements involved in motion control is greatly simplified and only the components pertaining to volitional motion of grasping and affordances are described using a bottom-up approach. Figure 2.4 is created to summarise the interactions between the components described in this section.

The most peripheral parts of the controller, which initiates intentional motor actions, are the alpha and gamma motor neurons. Such neurons are used to control every muscle of the body and they are used in similar ways across the organism [40]. The neurons

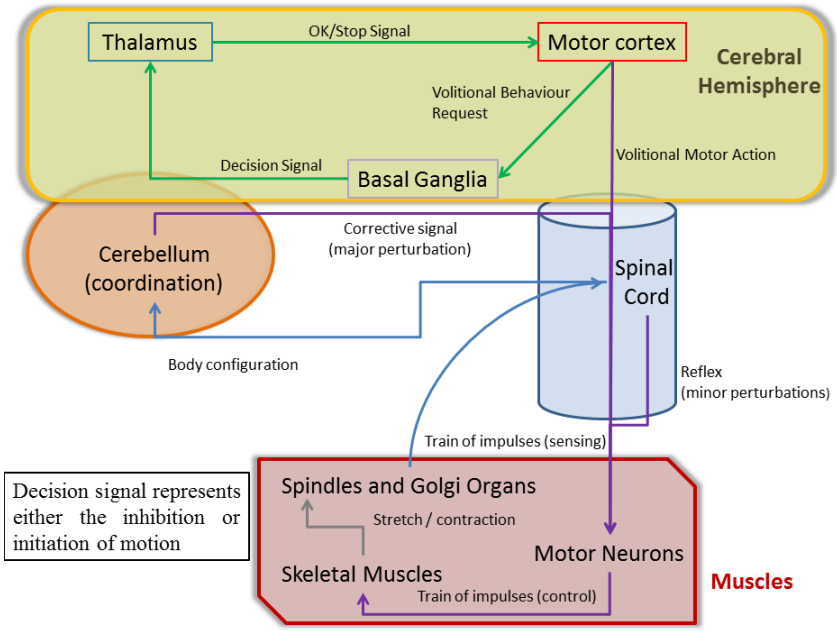


Figure 2.4: Summary of the elements involved in the human action loop. Purple arrows represent motor commands, blue arrows represent sensory information and green arrows represent neural processing internal to the cortex. A major perturbation is an unexpected event that requires a large motor correction such as avoiding falling.

work in pair by acting electrically through a train of impulses of variable frequency on different parts of the muscle [41]. For example, the higher the frequency of alpha motor neuron impulses the more the muscle will contract [42]. The motor neurons extend from the motor cortex directly to the controlled muscle via the spinal cord. The spinal cord micromanages a complex motion and locally corrects for unexpected situations by triggering reflexes on the motor neurons [43]. It is believed [44] that the motor neurons acting on the muscles of the hand do not work independently from each other, but rather in a synergistic way. In other words, the CNS controls groups of muscles that operate on different joints at the same time rather than independently.

At all times, the CNS is aware of the displacement and tension of each muscle, as numerous proprioceptive sensors are distributed in the body. The sensors used to detect

the stretch of the muscles are located in the muscle spindles. The spindles are fusiform sensory receptors located in the thick part of the muscle. Neurons afferent to the brain are connected to the spindles and are activated during or after stretching a muscle. The afferents fire a train of impulses of higher frequency while the muscle is actively changing length, or after the muscle is stretched to its new length [45]. A typical use of the spindle is to detect whether a movement failed or to trigger reflex motions. The muscle tension is measured via the Golgi organ [46]. This structure is located between the end of the muscle and the beginning of the tendon and it is sensitive to muscle tension. The neuron afferents of the Golgi organ are very sensitive. They activate when the organ is stretched as the muscle contracts, and have a rate of firing proportional to the applied strength. A coordinated use of the Golgi organs and the spindles is required to detect and deal with unexpected situations by triggering reflexes via the motor neurons [41]. For example, this happens when a lifted object is heavier than expected and a quick motor correction signal is required. Such operations are performed locally in the spinal cord.

Near the occipital area of the brain lies the cerebellum. One of its roles in motor control of grasping is to perform motion planning and to coordinate spatially and temporally the plant during motion execution [47]. For example, it is acknowledged that the cerebellum supervises reaching execution and digits coordination when preshaping the hand. The cerebellum is just 10% of the volume of the CNS, but it contains more than the half of the total number of neurons [48]. The cerebellum is linked to the thalamus, which is the centre of all the sensory input. Additionally, it is believed that error based learning of motor actions and grasping skills is implemented in the cerebellum [48]. In this area, sensory feedback from the thalamus is integrated with executive commands and, if an error is detected, online motor corrections take place [49]. It is also believed that learned actions are retained in the cerebellum and automatically recalled when needed once the new action is well understood by the subject [50]. In general, learning and adaptation

is a continuous process in the CNS [51], which initiates at birth and shapes the neural structure of an healthy brain until the end of life [52].

Within the cerebral hemisphere (the brain), the motor cortex is responsible for the initiation of voluntary movements, such as grasping a cup of coffee. The cortex does not micromanage muscle motion; it rather encodes behaviours that can be initiated voluntarily or as an emotional reaction [53]. As it is difficult to precisely analyse cortical circuits directly, there is not complete agreement on some mechanisms of the motor cortex. The motor cortex can be divided in two areas: the primary motor cortex and the premotor cortex. The role of the latter is currently under discussion, but it is believed that it is related to initiating body or facial movements as a reaction to an emotion, to control focus and to initiate movements and social communication as a response to sensory stimuli [53]. In the primary motor cortex, the intention of initiating and directing a voluntary action is elaborated [54], such as deciding to reach to grasp an orange rather than an apple. This part of the cortex also defines the direction, amplitude and force of behaviour [55]. As such, the primary cortex initiates the motor control loop, which allows the body to implement a volitional action. Different and contiguous areas of the primary cortex are controlling different parts of the body, and the finer precision a body part requires, the larger the area dedicated [56]. It was understood that the control of fingers and skilled manual behaviours occupy a significant part of the cortex.

Once a volitional behaviour is decided in the primary motor cortex, Basal Ganglia are activated prior initiating or stopping a motion. In the context of motor control, the Basal Ganglia suppress motions that are not required to accomplish the requested behaviour, such as hand tremors, and they have a role in initiating and terminating correctly a motor sequence [57,58]. They also influence the mood and cognitive functions [48].

It is believed that object and grasp affordances are elaborated in the cortex, although its exact mechanism is still debated. To initiate the process of performing a voluntary

afforded action, the perception of a target object generates neural activity in different areas of the cortex, processing the past experience and other environmental constraints [59–61]. The result of this activity defines a suitable motor execution specifying the grip type and the grasp phases, and the grasp affordance [59]. Perception can recall many different grasp affordances, since an object can be grasped in multiple ways. The most suitable grasp affordance is then selected through a competition [7, 60] biased by the available higher cognitive information, such as rewards, short term memory and planning.

2.3 Robotic Systems for Grasping and Reaching Objects

2.3.1 The Plant: Multifingered Robotic Hands

In this section, robotic hands are described in terms of mechanical, control and kinematics characteristics and compared to the equivalent characteristics of human hands. Robotic multifingered hands can be broadly divided in three categories: fully actuated, under-actuated and actuated/under-actuated hybrids. Fully actuated hands were the first to be researched [3] for robotics manipulation, and have all the joints in their kinematics controlled independently. Hands with hybrid actuation superseded fully actuated hands as only part of the hand kinematics was directly controllable while the other joints, such as the distal and medial phalanx joints, were mechanically coupled [62–65]. In under-actuated hands, a few variables, each representing a posture, are linearly combined and operate on all the joints to emulate the synergistic grasping, such as the one observed in humans [44]. First under-actuated, or synergistically actuated, hands were introduced in the late 2000s [66]. The actuation constrains of the joints, derived from processed human trials, are implemented in hardware. The actuation is posture specific and cannot be modified after manufacturing. A comparison between fully actuated and

under-actuated mechanisms is shown in Figure 2.5.

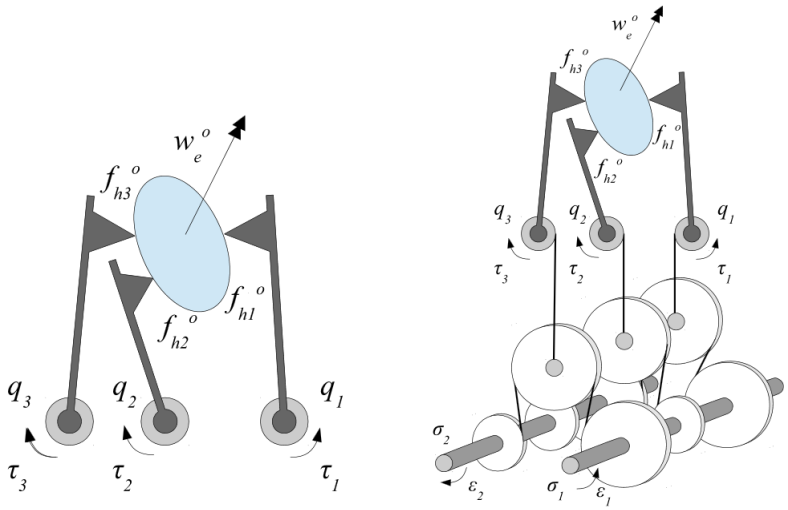


Figure 2.5: Comparison between fully actuated (a) and under-actuated joints (b). Fully actuated joints require the control of all the joint configurations q and all the torques τ. Under-actuated joints require the control of only the synergy configurations σ and the forces η as the motion is propagated to the joints. f are the forces applied to the object and w is the external wrench. Images adapted from [67]

.

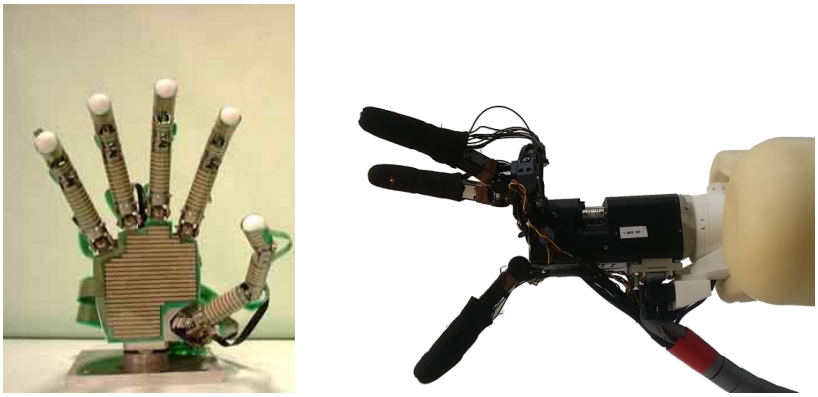
Although under-actuated hands often have a humanoid kinematic model, robotic multifingered hands do not have this requirement for grasping. The first fully actuated robotic hand prototypes had just three [23] or four fingers [68]. Later designs feature all the five fingers [63] or non-anthropomorphic kinematics [69]. State-of-the-art humanoid hands can have a large number of control variables [70], or as little as nine variables [62] and have hybrid actuation. Non-antropomorphic hands, instead, rely on different principles for grasping, such as a foldable palm [71] or an passive grasping envelopment [72], and it is debated whether a different kinematics gives an advantage on grasping and using everyday objects. Although the recent anthropomorphic hands closely resemble a human hand, their kinematic model is extremely different, and generally varies from a minimum of 15 [73] to a maximum of 25 [74] directly and indirectly actuated joints.

Under-actuated hands often have a similar kinematic model to hybrid actuated hands, and can be approximately 10 times cheaper than a hybrid actuated hand with similar kinematics [1]. The main difference is that the joints are not directly controllable. The rationale behind the used number of joints is rarely justified. In addition, the impact of chosen kinematic model on the manipulation potential of a gripper is not quantified, especially in relation to the kinematics of the thumb [75]. As the thumb is believed to be the main contributor to the dexterity of the human hand [35], such kinematic mismatch between the human and state-of-the-art robotic hands is a limitation. Indeed, this reduces the possible space of configuration of the fingers and, consequently, the variety of postures available for different manipulation tasks [8].

Traditionally, robotic hands are manufactured using hard materials such as plastic [62] and metal [76]. One of the most popular actuators are Direct Current (DC) motors, embedded either in the arm [62] or directly in the hand [71]. and pulleys, which are not as elastic as human tendons. Disadvantage of those common manufacturing choices is that hard materials make the hand non-deformable and the fingers not very compliant. Lack of deformation and reduced compliance limit the adaptability of the fingers and of the outer structure of the hand to the shape of the object. This reduces the friction created by a deformable padding and the ability to envelop precisely and object within fingers, which are mechanisms intensively used by human tendons and dermis when grasping. For these reasons, recent efforts attempted to use elastic and deformable components [77], variable stiffness [78] and pneumatic actuation [79,80] and soft hand designs [81]. Although the choice of hard materials limits the flexibility of robotic hands, state of the art manufacturing can achieve some degree of dexterity [82].

The human hand is one of the most sensorised parts of the body [85] and its sensory stimuli is very important for perceiving the environment, especially if blindfolded [86]. In

[1] Shadow Dexterous Hand



(a) Whole surface tactile sensors (b) Force, torque and proximity sensors

Figure 2.6: Example of different sensorised hands. The Gifu hand III (a) has tactile sensors on the whole surface of the hand [83]. Torque, force and proximity sensors [84] can be applied to the KCL metamorphic hand (b), which has a non-humanoid kinematics [71].

an attempt to emulate this principle, different types and number of sensors are installed on many robotic hands. Most hands, such as the KCL metamorphic hand or the iCub hand, are able to obtain proprioceptive feedback extrapolated from the motor encoders or the current applied, providing joint information comparable to human motor afferents. Such type of sensing is very widespread and simple to use for motor control [87], but it does not provide information on the mechanical proprieties of the object. Additionally, a continuously applied high current can damage the motors, while human fingers are not damaged from if continuously gripping with strength. For this reason, various tactile and force sensors are increasingly used for robotic manipulation. Different implementations of tactile sensors can detect the contact point with the objects and they can be installed on the fingertips [88] or the full fingers and palm [83,89]. Force and torque sensors [90] are used to measure the magnitude of the contact. Example of those technologies are shown in Figure 2.6. Some common limitations of those sensors are the cost, the maximum load, the reliability of the measurements, and the limited domain of use. However it is expected that their role will be more and more prominent to support

more advanced control mechanisms. Other human sensory perceptions not yet much considered in robotic manipulation are heat and pain. While a defence mechanism such as pain might not be needed in robotics, temperature can be useful to guide grasping as it can differentiate a living organism from an object, for example. Finally, non-biologically inspired types of perception, such as object proximity, have been already integrated in grasping system [91]. Good results have been achieved when different types of sensory information are integrated [84,92].

To summarise, there are a lot of kinematic representations of the human hand but it has not been established which feature is most important to render in such representations. Fully, hybrid or under-actuated hands give more or less freedom of movement to the fingers, at the expenses of additional complexity and cost. Traditional manufacturing and actuation technology lack of the intrinsic compliance of human hands, therefore alternatives which use soft material and different types of actuators. Finally, artificial sensors do not have yet the same level of sophistication or density of biological sensors, but it is possible to compensate for this by integrating in the hand hardware other sensing capabilities not used in nature such as vision (*camera-in-hand*), ultrasounds or proximity.

2.3.2 The Controller: Optimisation and Learning of Grasping and Reaching

The problem of robotic hand control for manipulation was initially approached in its closed form in the early 1980s. One of the first results was the mathematical formulation of the hand Jacobian [93]. It describes a set of equations that link the hand joints' speed and torque to the position and force of respective fingers. Grasping was performed by matching the produced forces or positions to the forces that are required to restrain the object. Force and form closures were formulated as metrics for assessing the grasp stability in closed form to compensate for the effect of friction [94]. At this point, a

(a) (b)

Figure 2.7: Intuitive illustration of force closure (a) and form closure (b). The first defines an object as constrained based on the forces applied, the second based on the points of contact.

successful grasp was performed by optimising the contact points or the applied forces on the object by the robotic hand [95]. An intuitive representation of the closures is shown in Figure 2.7. One disadvantage of such approach is that the computation of optimal contacts is a non-trivial and computationally expensive task, which requires knowledge of the object geometry and position. Later studies demonstrated that it is possible to optimise those calculations [96]. However, the mathematical approach still requires a re-computation of the optimal contacts for each new object that is manipulated. Thus, it is limiting the use of such technique in unknown and unstructured environments such as a household or a nursery. As it can be seen, the traditional mathematical approach is fundamentally different from the biological approach observed in humans and it is generally employed on fully actuated hands.

In an attempt to reduce the amount of required computation and the dependency on the specific object's characteristics, under-actuated hands were introduced [97]. Those

hands share a similar kinematic model of fully actuated hands but with different joint control and are briefly discussed in Section 2.3.1. In one of the first results [66], a small number of variables was used to control the motion of a robotic hand with complex humanoid kinematics. This type of control represents the translation of the grasping synergies observed in the CNS and the hand muscoloskeletal system of humans to robotics [44]. Subsequent works extended the original concept by taking into account the forces generated by the fingers [98] and the possibility to adapt the grip to the shape of the object [99]. The theoretical framework for assessing the quality of grasp for under-actuated robotic hands is not as mature as for the traditional approach, but recent works are closing the gap in this direction [100, 101]. Under-actuated hands are easier to control and can produce reliable grasps. However, the flexibility of such hands is limited by the hardware. Indeed, once a set of synergies is implemented, the hand and its control strategy cannot be changed and it is not easy to alter the type or number of synergies or to switch to a different control strategy in real time. While this choice might be suitable for prosthetics, where simplicity of control is fundamental [102], robotic applications can be limited by this choice, as different tools might require a more or less specific set of synergies or a control switch during execution. It is, however, possible to implement a synergistic control in software [103, 104] without losing the flexibility of a fully actuated hand.

Learning by demonstration is a popular technique used to simplify the tuning and implementation of reliable models for grasping and other structured actions. Learning by demonstration, also known as imitation learning, is an alternative to adaptation via trial-and-errors. Adaptation methods are adjusting a model incrementally during its execution using strategies, such as adaptive control [105] or reinforcement learning [106], compensating for the errors during execution. A disadvantage of those approaches is that a large number of trials are required and it might be difficult to achieve convergence [107, 108]. Indeed, finding the optimal parameters in the model's search space

requires long time. Therefore, it is impractical to realise such algorithms on a real robot that performs advanced manipulation tasks. Learning by demonstration, instead, allows adjusting or deriving a model of an action from data collected from demonstrations of the required task. The optimal parameters, or the model itself, are directly derived from the processed data of the demonstrations. For example, learning is often a starting point for collecting the data used for the set of grasping synergies to be used in the hand hardware [109]. Additionally, learning is also used to tune or adapt motor models that realise a specific skill or an action [110], such as pouring a liquid in a container [111]. The learning data can be obtained with the following methods: a) collection in the robot's extrinsic joint space, by employing the robot's sensors or external equipment to observe the actions of a teacher; b) collection in the robot's intrinsic joint space, by capturing the joint data during teleoperation or while teaching kinaesthetically the motion [112]. However, regardless of the learning approach, the first challenge that needs to be addressed is to solve the correspondence problem [113]. The correspondence problem states that, since the kinematics of the teacher and the learner are different, it is first required to map the joint space of the teacher into the joint space of the learner in order to ensure meaningful teaching. If the teaching is performed in the robot's intrinsic space, the problem is automatically addressed. However some teaching means, such as teleoperation, might not be suitable for learning manipulation skills due to the complexity of controlling an articulated arm and a multifingered hand [114]. Performing a demonstration in the robot's extrinsic joint space is easier and more intuitive, particularly for teachers unfamiliar with robotics (employees in the agricultural and service sectors). However, the correspondence problem is more difficult to address for complex kinematics such as robotic hands. Several attempts were performed to assess how well a robotic hand maps into a human hand [9,115]. However, it is still difficult to solve the correspondence problem entirely due to the simplifications and disagreements in the human hand kinematic models. As a result, learning by demonstration tends

to be successfully employed in restricted domains, with a predefined number of objects or specific motion strategies [116]. In addition, it is hard to scale the technique over a large number of objects or a more flexible motion control. Although learning from demonstrations and from trial and errors are separately well studied, the two approaches are rarely combined in robotics. For example, results on the mechanism of learning of the human cerebellum, which combines reinforcement learning through practice with taught demonstrations of a teacher, are rarely used as source of inspiration for robotic learning algorithms. Anyway, it is possible to extract the core features of human motor behaviour from experiments, and to apply only the relevant biological principles to robotic motor control. For example, in [117], the authors are modelling human touch strategies of soft objects. The same model was later implemented on a robotic platform [91] with good results. Another example is shown in [118]. The authors perform human experiments of pick and place, grasping with sensory constraints, to identify the conditions that favour an action plan over another. The model that defines the conditions and the plans is general enough to be transferred to a robot with adequate sensing capabilities for grasping and reaching.

In order to execute a grasping or manipulation action, a motor sequence should be generated. Such plan guides the motor control of the end effector and the arm towards the target. Its functionalities can be compared to the ones of the cerebellum or spinal cord during human motion execution, depending on the level of complexity required. Some controllers are able to guide the manipulandum respecting specific constraints [119] or reactively avoiding obstacles [120], others are more closely replicating the neural mechanism of human reaching [121] or rely on techniques extrapolated from the artificial intelligence planning community [122]. Regardless of the mechanism used, planning of the reaching action is rarely combined with grasp planning, and those two aspects of the motion are generally studied separately.

In an attempt to create a more flexible interaction with different objects, the concepts of object and grasp affordances were introduced in robotics. In both cases, the aim is to give robots the ability to handle alternatives when executing a task. Object affordances can be used to introduce flexibility in high order plans, such as an assembly sequence, when presence of uncertainties requires a replan. For example, the motion plan of a mobile robot might change based on the path suggested by different unseen elements of the environment [123], or the order of execution of a complex task might change based on the available objects [124, 125]. Grasp affordances, instead, identify different approaches to manipulate an object trying to fulfil a criteria [10], as the same object often allows to be grasped in very different ways. Often affordances are learned from demonstrations [126, 127] since the sequence of motor actions might be complex. Commonly, the methodology followed to translate affordances in robotics is comprehensive, as different problems, such as object perception, motor control and skill learning, are addressed by a single monolithic component. As a result, the existing approaches are very complex and specific, and do not allow the flexibility of object and grasp affordances to be exploited beyond the domain of certain application [128]. Conversely, the human planning and control system is specialised, hierarchical and general purpose and approaches inspired by such organisation are not frequent.

As the motor execution of complex sequences of varied actions in robotics is still considered to be very challenging, the creation and inhibition of composite plans of high level actions elaborated to fulfil a high level goal is not be discussed in this thesis. Such functions, performed by the basal ganglia and the motor cortex in humans, pertain more to the domain of artificial intelligence [129] and pure planning [130] which generally does not deal with the complexities introduced by the interaction with real world objects.

2.4 Conclusions: Requirements for Flexible Manipulation

The common characteristic of the above-mentioned studies of robotic manipulation is the focus on a specific manipulation task. This makes it difficult to formulate a general approach for grasp control of unperceived human-made objects. In this respect, the study of the generality of human manipulation skills can provide guidelines for future robotic implementations. On the other hand, it is important not to overestimate the contribution that studying human manipulation can give. Human studies shall be used to understand the principles that lead manipulation and translate the functionality rather than copying the mechanism. Replicating the structure of human neural circuitry to control grasping and learning [131] or shadowing the exact morphology of the human hand [132] might create solutions, which are domain specific and expensive to maintain or manufacture.

Another requirement to fulfil for robot manipulation is to reduce the dependency from the specific domain of application. If an unseen object is presented to humans for the first time, it is infrequent for them to handle it incorrectly using an inadequate grasp affordance. The amount of objects that are manipulated daily by humans is very large and continuously changing from day to day, making it extremely difficult to recalculate and to store the optimal grasping strategies or object models.

The enable a robot to interact with a large set of different objects, it should be assessed how different grasp affordances influence the performance of the final result of an action, rather than the fine details of a grasp posture or the object characteristics. The motivation is that most of the times grasping is a prerequisite for an action to take place for achieving a result. As long as the desired outcome is obtained, the details on the grasp affordance and object used are not as important. Also, there are much less actions to do than objects to grasp, so the variability and size of the search space is lower. In this

way the grasping problem is simplified, as it is not required to find the best possible grasp but a grasp sufficiently good to achieve the desired result. As grasp control would allow more flexibility, reaching can compensate for the limitations of robotic grasping and facilitate the use of different postures as multiple solutions, rather than a single best, are now viable. As such, grasping and reaching shall be studied together and considered as a part of the same problem when modelling grasp affordance controllers.

Traditionally, the approaches used in modelling grasp affordances combine different aspects of the problem. The perception, learning, decision-making, motor planning and execution processes are often addressed in the same algorithm in an unstructured way, making it difficult to perform changes, such as adding novel objects or behaviours. Although the human CNS performs all the above processes at the same time, it does show some sort of hierarchy and organisation. It is, therefore, preferable to similarly impose a modular separation of the different aspects of grasp affordance. This approach would enable to customise, adapt and tune the system to meet the performance criteria required for a given task. For example, the knowledge about an object location and its shape is enough to achieve a successful grasp, but this information does not need to be obtained from a visual perception of the environment only. In a modular affordance framework, the type of perception (e.g. visual or haptic) is not important as long as it can provide the information needed to perform the task within the expected efficiency.

To summarise the above, a list of requirements that robotic manipulator controllers shall address to increase their flexibility for different domains is proposed:

1. The fundamental principles, rather than the detailed mechanism, of human manipulation skills shall be studied and applied to robotic systems;

2. Performing an intended action effectively is the target of grasping, and the effect of the action given from a given posture shall be assessed;

3. Grasping and reaching for an object compensate each other and their interactions shall be analysed, but modelled independently;

4. Grasp affordances should be modelled in a modular framework, where different aspects of the problem can be addressed independently.

This thesis explores the above requirements from human studies (Chaper 3 and 4) and outlines the principles of a modular framework for grasp affordances (Chapter 5) and a biologically inspired grasping controller (Chapter 6).

To conclude, it is important to note that an object requires to be grasped before being used. Hence, it is first required to resolve the problem of general purpose grasping before addressing more complex problems, such as object affordances and high level planning, that abstract from the specifics of the action performed. Research in these directions is very useful, but it is not possible to fully exploit the results while robots are still unable to grasp different unspecified objects for a specific action. For this reason, this thesis addresses the problem of grasp affordances only. It is believed that to address high level reasoning problems, such as object affordance, it is first required to consolidate the modelling, learning and execution lower motor control skills.

Chapter 3

Modelling the Approach Phase of Grasp Affordances

3.1 Introduction

This chapter analyses patterns of approaching to grasp observed in humans and proposes suitable models of the observed behaviours. In robotics literature, grasp affordances are modelled as an aggregated system that comprises different aspects from perception to motor control. In this thesis instead, a grasp affordance is considered as a modular part of a larger system. The purpose of this chapter is to propose a structure of grasp affordances, in the form of mathematical models learned from human trials.

A grasp affordance action can be divided in two stages: approaching an object and grasping it. In this chapter, the term *approaching* is derived from the definition of grasp affordance of Section 2.1 and refers to the act of reaching an object with the intention of using it to perform an action. This is different from the term *reaching* which implies a displacement of the hand to a defined position; for instance when touching a surface,

pressing a button or positioning an industrial robotic end-effector for soldering. It is important to stress this difference, as an open-loop reaching or grasping action, without the intention to use the object, is not sufficient to obtain a grasp affordance. An open-loop action as described above gives no guidance in selecting the most appropriate motion.

To grasp an object, a robot requires first to reach it. Controllers for reaching have better overall performance comapred to grasping controllers and are commonly used in industrial set-ups. Direct reaching is a mostly a solved problem in robotics [120], while constrained reaching, such as obstacle avoidance [133] trajectory following [119] within a strict time limit [134], are still considered as research topics. The combination of reaching and grasping controllers in robotics, instead, are not investigated as much. Reaching and grasping are often considered as separate problems, although some studies from neuroscience of grasping suggest the contrary [135, 136]. Therefore, it is worth to study the interactions and relationship of reaching and grasping to improve robotic control systems. A good reaching might compensate for a bad grasping, or a different reaching trajectory might be required for a different grasp posture.

There are various studies that combine reaching and grasping in order to obtain a better understanding of the environment, to learn how to use an object, or to guide the hand effectively. Although those studies do not openly discuss reaching and grasping interactions, they do take them into account and are discussed here. Often, perception and learning aspects are included in the analysis of the combination of reaching and grasping. In the interactive perception technique, the robot builds a representation of an object by interacting with it and observing the outcome of its actions [137]. An application of grasping with a further intention to use an object can be seen in [138]. In this work the authors also employ interactive perception to teach a robot to use tools. Additionally, this technique, combined with Gaussian processes, can be used to

interactively determine how to fold laundry [139]. The combination of manipulation and reaching in interactive perception is used to improve the knowledge of the environment, and to understand how to interact with it. Therefore, perception and learning are fundamental components of this technique.

Another comparable approach is active vision. This methodology originally addresses complex computer vision problems by changing the view point of the camera [140]. Such technique can be used to optimise the number of processed frames needed to execute a grasp [141], or to generate grasping points on-line to guide visual servoing [142]. Hence, the interaction between vision and reaching is used to guide grasping. However, as the end-effector is mounted on the same arm as the camera, grasping is influenced by reaching. The above approaches study the interaction between reaching and grasping but do not openly target the actions as interdependent processes. In such way, it is difficult to understand the phenomenon in depth and scale it for different objects and domains (areas of application). As such, the main disadvantage is that those approaches are tailored to the specific problem. Additionally, an intense use of learning, required by interactive perception, often involves long on-line training for parameter tuning or model definition.

This chapter establishes a first step towards a modular grasp affordance system. Different parts of the motion, such as approaching and grasping, can still be interrelated without losing their generality. The methodology pursued in this chapter analyses of the approaching part only of a grasp affordance from human demonstrations. Further on, a model that describes the general pattern is provided.

This chapter addresses the question of whether or not, a general, object independent model of the approaching part of a grasp affordance can be derived from human demonstrations. In addition, it is studied whether a grasp affordance approach motion is a planned strategy or it is a set of reactive adjustments performed during the execution.

The contributions of this chapter are:

1. The interaction between reaching and grasping is characterised by analysing human grasping experiments, and expressed in terms of hand to object distance.

2. The motion pattern structure is defined in terms of approach distance and speed of fingers displacement.

3. A set of object-independent most reliable models is derived from the data to describe a general, object independent pattern of approach to grasp.

4. Alternative simpler and approximated models of the motion are derived to allow a comparison in a robotic set-up.

5. The motion is characterised in terms of orientation and distance with respect to the object.

Such interpretation of grasp affordances helps shifting the attention away from the geometrical details of a specific object. Thanks to the generality of the models, it gains more importance the effectiveness of an action execution, giving a grasp posture, reducing the dependence on the specific domain of application.

The rest of the chapter is organised as follows: in Section 3.2 the experimental data and data preprocessing methodology are discussed. Section 3.3 presents the results of the data analysis and describes the phases of the approach distance. In Section 3.4 a set of models for the approach distance and orientation data is presented. Section 3.5 is the discussion and Section 3.6 draws out the conclusions.

3.2 Methodology

3.2.1 Experimental Protocol

The aim of this chapter is to understand and define a general pattern in human approach to grasp for performing a specific action. Such pattern should be independent from the object grasped, but it can be dependent on the task performed. It is important to underline that a specific action has to be executed, as this constraints the list of possible strategies and postures to the ones required for the task. In case an action is not defined, there is an open scenario where any approach to grasp strategy is acceptable and it is not possible to discriminate the most appropriate strategy and grasp affordance. The assumption is that part of the approaching action should be stereotypical and another part should be dependent on the specific object properties. This is motivated by the fact that humans are able to grasp most objects intuitively, but are not always able to proficiently use them without prior knowledge. For example, it is trivial to grasp a fencing handle, but it requires a lot of training to grip the handle for successfully keeping the guard. The action selected for the task performed in the experiments of this chapter is hammering on a point. This action was selected because it can be easily generalised to similar actions, such as insertion or pressing. Additionally, hammering is one of the first actions ever learned by infants [143], and it was used by prehistoric humans for crafting the Oldowan stone tools [17]. Hence, this action can also be used for other simple scenarios such as basic crafting.

Approach to grasp and object motion data were collected from human trials for this study. For this purpose, it was important to track the hand, wrist and fingers motion in order to record trajectories and finger postures during a trial. Also, the object position and orientation are tracked through the whole experiment. This is required in order to highlight the overall grasp affordance decision process by relating the object position

and orientation to the grasping motion of the human participant, instead of processing the two independently.

Figure 3.1 shows the complete set up of the experiments. A system of four motion tracking cameras (Vicon Bonita) was used to record the object position with a frequency of 100 Hz. Additionally, the position and orientation of the subject's arm and hand was tracked using wearable fibre optic motion capture (MoCap) system (Measurand ShapeHandPlus) with a frequency of 77 Hz. To aid in the analysis, the frequencies of the two devices were adjusted to match 100 Hz using linear interpolation.

Participants were seated on a chair in front of a table, and asked to wear the MoCap system on their dominant side. The table was placed in the centre of the field of view of the four cameras, and it was covered with a black cloth to eliminate possible reflection from artificial light. The room was lit using artificial light only, and the illumination was kept constant throughout a capture session.

A set of eight everyday-use objects of different shape and weight, was used for the experiments: a plastic ball, a paper coffee cup, a card box, a phone headset, a CD keep case, an hard-cover book, a computer mouse and a hammer. Table 3.1 shows the properties of the objects. The purpose of the set up was to encourage the selection non-trivial hammering surfaces and unusual approach to grasp strategies, so that different motion sequences would be explored. In this way it is possible to define a general object-independent grasping pattern given the variability of approaches.

Nine subjects, seven males and two females, performed the experiments. Participants had no history of previous motor impairments and they were right handed. The mean hand width was 79.7 mm, the mean hand length was 189.3 mm. The measurements were performed using the methodology proposed in [27] based on the hand breadth and length from digitizer criteria. During experiments, the objects were selected in random

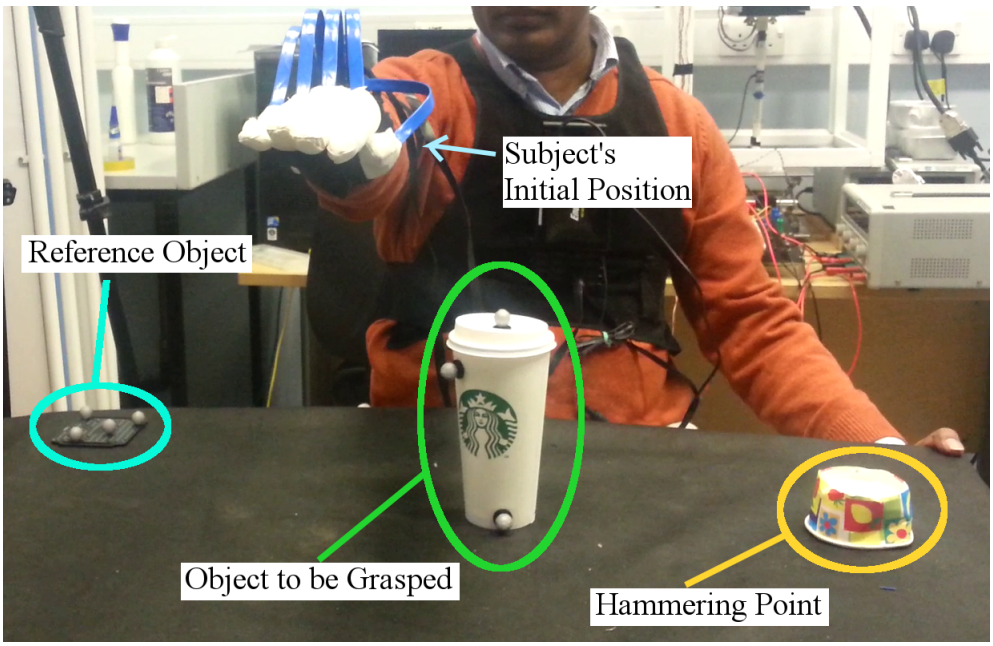


Figure 3.1: Illustration of the complete experimental set-up.

order and, when applicable, random orientation - the longitudinal axis of the object was either parallel or orthogonal to the table edge. Each subject performed the experiments using the objects placed in two different orientations, when possible. The cup and the ball do not have a unique orientation due to their circular base. Each approach to grasp experiment was repeated two times. In total, 28 demonstrations were collected for each participant. Experiments were approved by the King's College London Ethical Committee, REC reference Number BDM/12/13-27.

At the beginning, a hammering point on the non-dominant side of a subject was marked on the table using a paper cup as a damping place-holder. The point was placed off the central line of the body to further encourage subjects to explore different approaching patterns. The point was selected so that it would be possible to approach, grasp and hammer without the need of bending or rotating the torso. A small platform with 5 trackers was used as a common reference point for the Vicon and MoCap systems , as it is shown in Figure 3.1. It was located close to the dominant side of the subject on

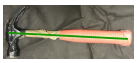
Object	Longitudinal Axis	Dimensions (cm)	Approximate Shape
Ball		⊘ 8	Round
Book		24.2x19.1x4	Rectangular, heavy
Box		72x20x20	Rectangular, light
CD Case		18.5x13.5x1.5	Flat
Cup		13.2 H, ⊘ 8.5	Cylindrical, wide
Hammer		33x12x2	Cylindrical, thin
Headset		21x7x4	Rectangular, thin
Computer Mouse		14.5x13.7x5	Oval

Table 3.1: List of objects used in the experimentation and their properties. The longitudinal axis is highlighted in green on each object's picture. The first dimension is along the longitudinal axis, the second is orthogonal to the axis and lying on the same plane. ⊘ stands for diameter, H stands for height.

the corner of the table. The position and orientation of the reference plate are fixed for the duration of the whole trial. The object to be grasped is positioned in front of the subject to allow comfortable approaching and grasping without the need to bend the torso. The participants were able to perform a direct approach motion and no obstacle was impeding their action when performing the experiment.

Subjects were shown a demonstration of the experimental protocol prior starting and given further clarifications if needed. At the beginning of each trial, subjects were asked to adopt the initial reference posture shown in Figure 3.2. Further on, subjects performed the experimental protocol as follows:

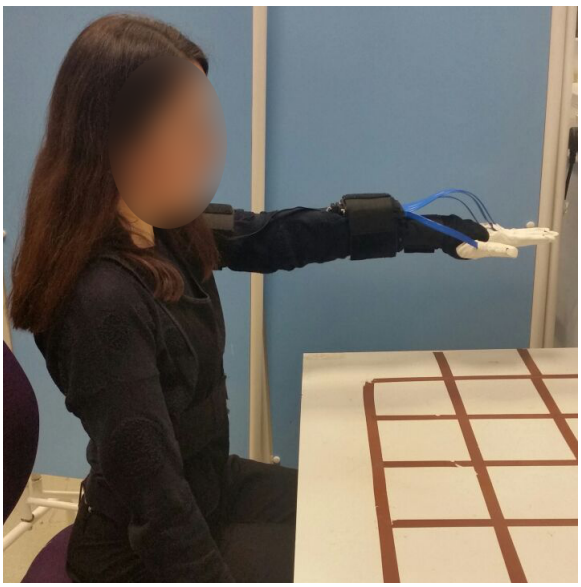


Figure 3.2: Initial reference posture: arm in straight position parallel to the ground and orthogonal to the chest, hand fully open and flat fingers.

1. Subjects covered the reference plate with their hand so that it is not visually tracked. Losing the tracking allows synchronising the starting point of both data streams.

2. Subjects returned their hand to the initial posture, so that the reference plate is tracked again (Figure 3.2).

3. Subjects approached the target object to grasp naturally. No constraints or suggestions were given regarding the most suitable grasp affordance.

4. Subjects hammered the grasped object on the marked area. Subjects were free to choose the hammering style or object's point of contact with the hammering area.

5. Subjects placed the object away and the data collection was stopped.

Although the whole action listed above was captured, only the motion between steps

3 and 4 was analysed in this thesis. Performing the complete action was required to ensure that subjects would perform a natural approaching to grasp pattern, so that any difference with plain reaching could be highlighted.

3.2.2 Data Processing

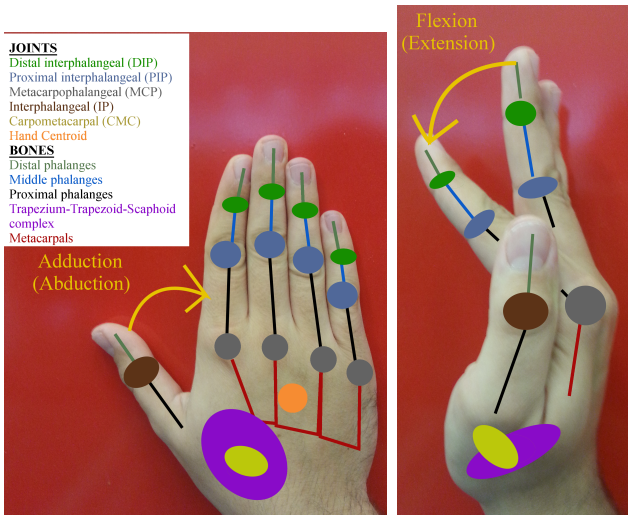
Data Collection Methodology

In the analysis the relationship between hand position and the object centroid is characterised. The centroid positions of the reference plate and the object were acquired directly from the visual tracking system with no need of further processing. The positions of every joint of the kinematic model of the arm were collected for every trial. The hand position is defined as the centroid between the wrist, middle and ring fingers metacarpophalangeal (MCP , see Figure 3.3(a)) joint positions . The positions were derived from the kinematic model using the MoCap toolbox for MATLAB [144].

Metacarpal, proximal and distal interphalangeal joint angles, and metacarpal adduction/abduction joint angles of the index, middle, ring and little fingers were also recorded (Figures 3.3(a) and 3.3(b)). In this study, only MCP flexion/extension joint motion of the fingers was analysed, as it has the greatest impact on the motion of the whole finger [145]. Further in the text, the MCP flexion/extension data is referred as metacarpal data.

General Data Processing Methodology

Visual tracking and wearable MoCap data were synchronised as part of the experimental protocol. The moment when the reference plate was covered by the hand of the subject was considered as the common starting point. The reference frame of the MoCap and



(a) Top view (b) Side view

Figure 3.3: Schematics of the hand bones and joints involved in the capture.

visual tracking data were transformed to the coordinate frame of the reference plate in order to allow comparisons in a common coordinate system.

The approaching to grasp part of the entire trial was isolated by inspecting the motion of the fingers and object. Specifically, the first displacements of the fingers marked the start of the analysed sequence. The end of the sequence was established at the moment when the object's centroid had a vertical speed higher than 3 mm/s. This speed value clearly indicates that the object is being lifted to be used. For this to happen, the object has to be secured in the hand first, hence it is a clear indication of a completed grasp action. Although the analysis focussed only on the approaching to grasp part, subjects were required to perform a complete and realistic hammering action. This was required to ensure that the approaching to grasp part of the motion would not be artefact for the experiment.

The normalised euclidean distance between centre of the palm and the object centroid was calculated, and used to quantify the relationship between the hand and object po-

sitions for the approach distance analysis. The normalised euclidean distance is defined as approach distance and it was calculated as follows:

$$a = \sqrt{\sum_{t=1}^{n} \left(\frac{O_d^t}{O_d^{mx}} - \frac{P_d^t}{P_d^{mx}} \right)^2} \tag{3.1}$$

Where a is the approach distance, O_d^t and P_d^t are the values of the object and palm respectively for dimension d observed at instant t. O_d^{mx} and P_d^{mx} are the maximum position values for dimension d of the object and palm. Finally dimension d refers to x, y and z axes.

The purpose of the normalisation is to transform the approach distance in a non-dimensional quantity for a clearer presentation. Since the data is already expressed in the same range, normalisation is not otherwise required. Hence, the normalised approach distance is a non-dimensional relationship between the hand and the object. It quantifies the distance between them at a given moment in time.

Data Processing Methodology for the Orientation of the Hand

The orientations of hand and object were expressed using unit quaternions. The hand rotation was extracted from the real part of the hand quaternion. The hand-object angular relationships were calculated from the axes of the hand and object quaternions. Such relationships define the zenith and azimuth approach angles of the hand to the object between the rotation axes of the two elements. Since the object is steady for the whole trial, it was possible to use it as a reference. The azimuth of the approach motion is the angle between the hand and the object axes of rotation on the horizontal plane of the table. The zenith of the approach motion is the angle between the two axes on

the vertical plane orthogonal to the table. Figure 3.4 describes all the calculations.

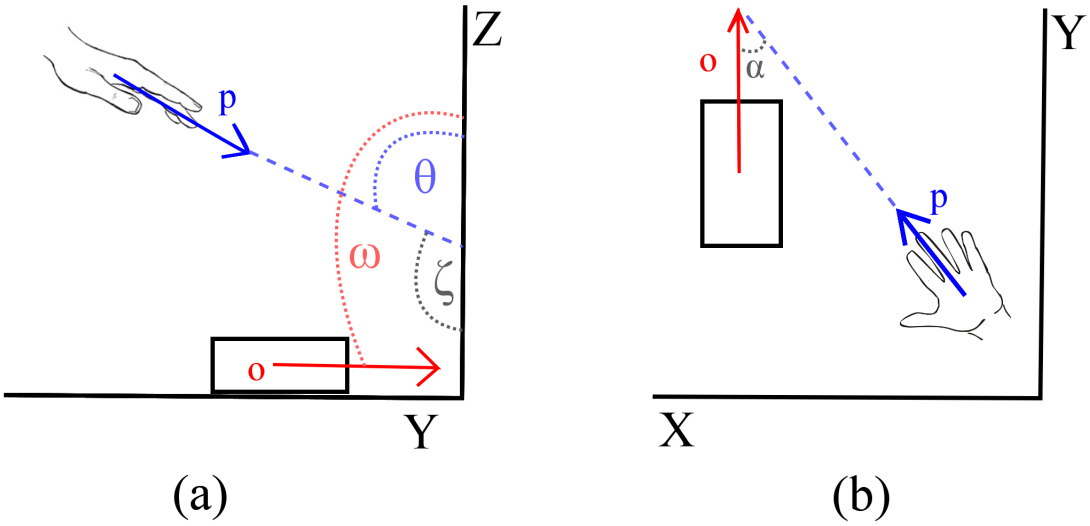


Figure 3.4: Geometrical description of zenith (a) and azimuth (b) angles of approach.

The azimuth was obtained from the dot product (projection) between the XY components (width and depth) of the palm and object quaternions as follows:

$$\alpha = \arccos\left(\frac{p_{q_{xy}} \cdot o_{q_{xy}}}{\|p_{q_{xy}}\|\|o_{q_{xy}}\|}\right) \tag{3.2}$$

Where $p_{q_{xy}}$ and $o_{q_{xy}}$ are the XY imaginary parts of the unit quaternions of the palm and the object and α is the angle between the two vectors (azimuth).

The zenith angle was derived from the difference of the direction cosine angles of the palm and the object with respect to the Z axis (height) as follows:

$$\theta = \arccos\left(\frac{p_{q_z}}{\|p_q\|}\right) \tag{3.3}$$

$$\omega = \arccos\left(\frac{o_{q_z}}{\|o_q\|}\right) \tag{3.4}$$

$$\zeta = |\theta - \omega| \tag{3.5}$$

Where p_{q_z} and o_{q_z} are the Z components of the imaginary part of the unit quaternions of the palm and the object respectively, p_q and o_q are the full imaginary parts of the palm and object unit quaternions, θ and ω are the angles between the Z reference axis and the palm and object unit quaternions imaginary parts and ζ is the zenith angle.

The approach positional distance and orientation data was derived two times, to obtain the speed and acceleration that were filtered with a moving average filter with span 7.

3.3 Results

3.3.1 Statistical Analysis

Statistical analysis of behavioural data was conducted to test whether factors such as the grasped object, the performing subject, or the specific execution influence the approach motion. Since a common object-independent approaching to grasp pattern is defined, statistical tests are required to verify whether every trial can be treated independently or all trials should be clustered and analysed together or in groups. Therefore, an Analysis of Variance (ANOVA) test was performed on the approaching to grasp data.

The statistical analysis was performed to understand whether specific features of approaching to grasp depend on the object, the specific trial or the performing subject.

The standard deviation of approach distance, azimuth, zenith and rotation were normalised and used as dependent variables to evaluate their relationship with the performing subject, the grasped object and the executed trial. It is expected to see no significant difference across trials if the standard deviation is similar. The independent variables were subjects, objects and trial sequence numbers for an object-subject combination. Three hypotheses were tested using a one way ANOVA test with one inter-cell degree of freedom and 230 within cell degrees of freedom ($F_{1,230}$). A hypothesis is considered significant if the Fisher's index (F) is bigger than F critical and the null hypothesis is rejected with 95% confidence level, which corresponds to a probability distribution (p) less than 0.05. The variances of the approach distance, the hand rotation and the zenith of approach did not show a significant dependence on the object being grasped ($F_{1,230} = 1.75,\ p = 0.19$, $F_{1,230} = 0.0012,\ p = 0.97$ and $F_{1,230} = 0.018,\ p = 0.89$ respectively) or the specific trial ($F_{1,230} = 0.04,\ p = 0.85$, $F_{1,230} = 0.0047,\ p = 0.82$ and $F_{1,230} = 1.11,\ p = 0.29$ respectively) but it did show a significant dependence on the performing subject ($F_{1,230} = 4.93,\ p < 0.001$, $F_{1,230} = 18.24,\ p < 0.001$ and $F_{1,230} = 4.05,\ p < 0.001$) respectively). Therefore, the performing subject is a determining variable of the structure of the approach motion for grasping. This result can be explained as every person performed the experiment at his or her own pace (speed). As some subjects were more careful or more confident, the speed of the execution was not consistent, although every object was approached to be grasped more or less in the same way.

The azimuth of approach did not show a significant dependence neither with the object being grasped ($F_{1,230} = 0.19,\ p = 0.66$)), nor the trial executed ($F_{1,230} = 0.07,\ p = 0.78$) nor the performing subject ($F_{1,230} = 1.095,\ p = 0.3$). This shows that the orientation of the hand on the planar surface the table is not discriminated by any external factor, suggesting that a common approach strategy exists across subjects and the same pattern is used when approaching different objects. This result can be explained as each subject

orients the hand in respect to the planar surface of the table in the same predefined way to complete a grasp of any object, while the other quantities might be more influenced by the subject's pace.

This result shows that the data should be grouped by subject for the analysis as the performing subject is a factor that influences most of the characteristics of the motion.

3.3.2 Characterisation of Approach Distance Patterns

In this section the analysis of the data is discussed and a common structure of the approach motion is formulated. The data is analysed by observing the individual patterns of motion of the approach distance speed, acceleration and fingers speed variability. As the purpose of this chapter is to analyse approach motions of the hand to the object, the speed of the approach distance is the fundamental quantity analysed. This value represents the rate of change of the distance between palm and object over time. The acceleration is considered in order to highlight changes in this fundamental quantity and to give structure to the motion. The hand position during approaching is considered, but not analysed here as it was found that it does not highlight well enough the dynamic changes involved during approaching. For similar reasons, the speed of the metacarpophalangeal (MCP) joint displacement of the index, middle, ring and little fingers are analysed. To better highlight the points where the MCP joints displace the most, the variance of the four fingers across the motion is examined. A variance higher than average, indicates a part of interest for the analysis since large MCP joint displacements imply that the hand is performing an activity such as preshaping.

Empirical comparisons of each trial showed that four phases can be discriminated, as shown in Figure 3.5. The first three phases represent the approach to grasp motion. The last phase comprises the final stage of grasping, when the object is firmly enveloped by

the fingers, and the beginning of the lifting motion of the object is performed prior to hammer. Each phase, except for the last, has its own characteristics which are common across all the trials.

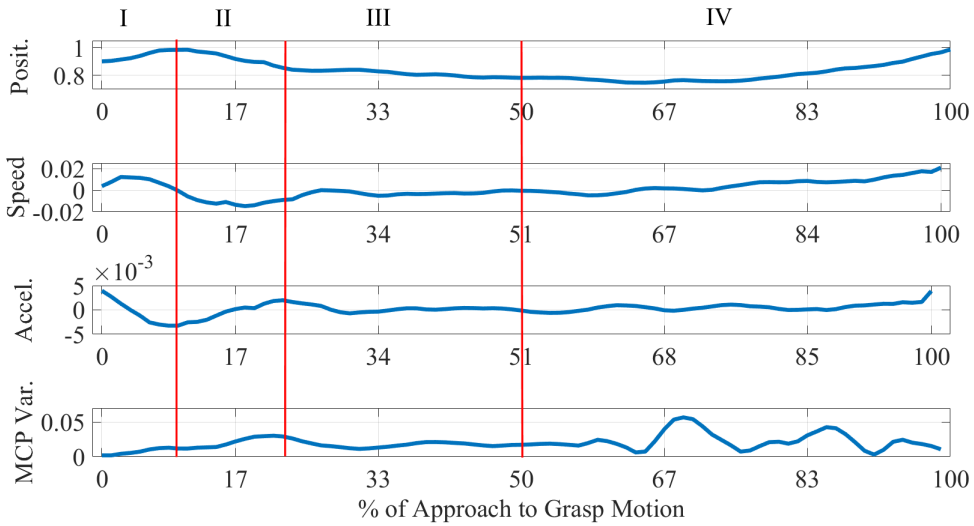


Figure 3.5: Sample approach to grasp trial. From top to bottom plot: hand position, speed, acceleration and variance of the four metacarpophalangeal joint speeds for the whole approach to grasp motion over time. The first data point corresponds to the moment the hand and the finger start to move, the last 17% of the motion shows the object being lifted for hammering as the speed of motion is ascending to the maximum. The Roman numbers identify the four phases.

To demonstrate that the data is similar across the dataset, a correlation analysis is performed on each phase of the segmented trials. The analysis is performed on positional data since it is the least processed data. The main issue to address is that subjects performed experiments at their own pace. Therefore, the length of a phase or of the whole experiment is influenced by external factors such as the subject's emotions (rush, boredom, etc.). For this reason, the four phases, shown in Figure 3.5, are analysed independently, and the duration of each phase for different trials is matched through interpolation. In this way the pattern structure within the phase is preserved. Each trial is segmented one by one according to the criteria defining each phase. The pairwise

correlation coefficients for the trials are calculated, and the overall median value of all the coefficients is considered. The correlation coefficients for the first three phases are 0.93, 0.99 and 0.97, respectively. This shows that the observed structure of the motion and the characteristics of each phase are common across all trials irrespective of the subject and the approached object; this matches with statistical tests. The fourth phase, instead, shows a median correlation coefficient of 0.16. This demonstrates that the phase is performed in different ways for each trial. This result is interesting since this phase shows a high MCP joint speed variability, as the fingers are finalising the grasp, and a low hand position correlation coefficient. This suggests that both finger and hand motions are important factors in grasping, and a separate analysis is required to better understand their relationship. Below, the four phases are discussed more in detail.

First Phase

In the first phase, the hand starts its approach motion to the object and the finger MCP joints just begin to move. This phase covers an average of $17.86\% \pm 6.92\%$ of the total motion sequence across subjects. The distinctive features of this phase are as follows:

- the hand speed increases to a peak and then starts decreasing;
- the hand decelerates abruptly until its global minimum;
- finger posture starts to shape from the initial steady flat hand configuration;

Although the MCP joints are moving, the variance of this motion does not significantly exceed the maximum mean peak in this phase that is 22% times smaller than the global mean maximum. This suggests that the fingers are displacing as the preshaping is just started. Most of the finger preshaping motion is performed in the next phase.

Another notable characteristic is that the speed profile is a bell-shaped curve resembling a Gaussian. This profile is distinctive for open loop reaching motions, as found in [146]. The main difference from that work is that, in this case, the bell-shaped profile terminates before the whole motion is completed, while in [146] the profile is extended until the end of the motion. This suggests that subjects treat differently an open loop reaching motion from a targeted approach motion.

Second Phase

In the second phase, most of finger preshaping motion is performed and the hand motion patterns undergo important changes in speed and acceleration. This phase covers as average $12.03\% \pm 2.10\%$ of the total motion across subjects. The following features characterise the second phase:

- the hand speed stays within its global minimum range;
- the hand acceleration increases until its first peak;
- most of the preshaping is performed, as fingers' MCP speed variability is increasing;

The MCP speed joint variance is above 72% of the total variability, suggesting that most of the preshaping is performed in this phase. Such increase of variability is not seen anywhere else than the fourth phase. This indicates that subjects select the finger posture to be used for grasping by the end of this phase.

Additionally, for each individual subject, the variability of the speed patterns undergoes a bell-shaped increase, underlining that the hand approach pattern is also adjusted in this phase. Since the MCP variability do not significantly increase until the last phase, it is likely that the subject adopts the actual hand approach pattern and the finger

posture to use for grasping. This suggests that the approach to grasp motion is decided and adjusted on the way rather than being preplanned. Section 3.4 discusses whether the adjustment of the approach grasp motion is reactive or intentional.

Third Phase

In the third phase, the distance between the hand and the object reduces until the approaching motion is terminated. In this phase the speed and acceleration of the hand motion reach a steady state. This phase covers as average $30.05\% \pm 3.75\%$ of the total motion across subjects. This phase is characterised as follows:

- the hand steadily increases its speed until settling down to 0 (± 0.001) mm/sec;
- the acceleration slowly converges to a steady state value of 0 (± 0.0001) mm/sec;
- the hand closes up the distance with the object to finalise the grasping, as the fingers' MCP joints speed variability change is minimal;

In this phase the finger MCP joint speed variance also greatly reduces until reaching a steady or null speed in some cases. This indicates that the implementation of the finger posture, selected in the previous phase, approaches its end until the fingers stop moving. This happens just before the actual grasp is performed when the fingers envelope the object.

The hand approach speed and acceleration also settle down to a more predefined pattern since the variability of those two quantities greatly reduces. It can also be observed that the speed pattern converges exponentially to a steady state value. Such change is observed in all trials, although the time required to reach the settling value might change. This confirms that in the third phase the approach pattern and finger posture strategies are implemented, as by the end of this phase the hand is steady and the

fingers are not displacing, as they are ready to grasp the object.

Fourth Phase

In the fourth phase the approaching motion is completed and the object is constrained in the hand to be lifted for the subsequent action - hammering. This phase covers on average $40.06\% \pm 7.18\%$ of the total motion across subjects. The common characteristic is that fingers' MCP speed variance is changing as the final enveloping and in-hand adjustments of the object are performed. In addition, the speed of the hand can show a sharp increase in the final part as the object is lifted, possibly to compensate for the object's weight or particular geometry. This phase is the only part of the motion that is different across trials, and no common features are identified in the hand motion patterns, as in some trials the speed was steady, and in others oscillatory components were observed.

This phase corresponds with the second part of our definition of grasp affordance, where a specific grasp posture is employed on a precise part of the object and the grasp is completed. As such, the detailed characterisation of this phase is beyond the scope of this chapter. A possible model of this phase is sketched in Section 3.4.

3.3.3 Characterisation of Orientation Patterns

Speed of the orientation data was also analysed to obtain a more complete picture of the approach to grasp motion. A sample trial showing the three parameters of orientation is shown in Figure 3.6. The azimuth, zenith and hand rotation were analysed empirically trial by trial, and it was concluded that the behaviour of those three quantities is similar to the step response of a second order system such as a spring-mass-damper. All the three quantities have a common phase when the speed increases until a peak,

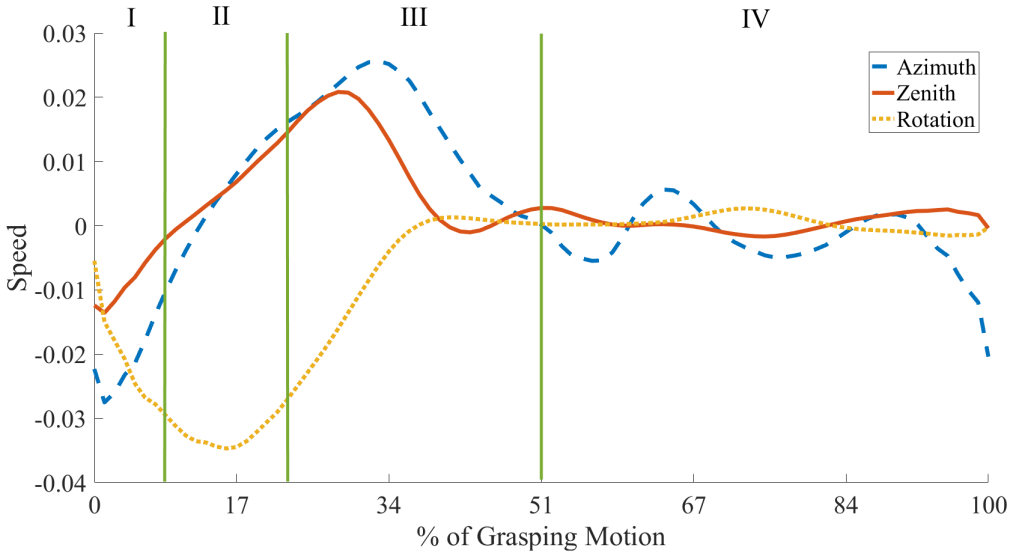


Figure 3.6: Evolution over time of the sample angles of approach to grasp trial. The Azimuth and Zenith angles of approach and the hand rotation of the same trial are overlaid and the four phases of approach distance reported for comparison. The first data point corresponds to the moment the hand and the finger start to move, the last 17% of the motion shows the object being lifted for hammering. The angles were measures as shown in Figure 3.4.

and a second phase where the speed decreases until convergence. The second phase is following a pattern very similar to the step response of a second order system. Although the orientation quantities have a behaviour similar to a spring-mass-damper system, they do have different convergence rates and reach the overshoot point at different moments. Section 3.4 investigates whether the orientation motions are reactive rather than intentionally planned.

3.4 Modelling of Approach to Grasp

3.4.1 Methodology

Different model types were fitted to characterise the speed of approach distance. The speed is analysed because the main features of human behaviour are captured in that domain, and velocity control is also commonly used for robotic control. Hence, a model for the speed of approach can greatly help to implement a robotic approaching controller.

The approach distance data was divided in the four phases, as described in Section 3.3.2. The regressive models of different orders, up to the fourth, were fitted to the positional data. The approach to grasp part of the grasp affordance is the principal phenomena modelled. However, models of the fourth phase are proposed to give better insights of the structure of this part of the motion. The maximum model order to fit was selected based on empirical evaluations.

The azimuth, zenith and hand rotation data were analysed as separate quantities. In this case, model fitting is used to verify the hypotheses whether the orientation pattern is a reactive or a planned motion. To verify it, the solutions to the differential equations describing the step response of first and second order models were fitted to the data. Indeed, first and second order models describe the dynamics of a reactive system, such a spring-mass-damper system. Hence, if the equation of such dynamics well describes an orientation quantity better than any other model, then, by analogy, that quantity has a reactive behaviour. The models used to verify the hypothesis of planned motion were fitted up to the sixth order, as the whole motion was more complex than the approach distance.

All the 224 trials were fitted independently to estimate the approach distance models. The quality of the fit was assessed via random sampling 10-fold cross-validation. 75%

of the dataset was used as training set and the other 25% was used as a test set. To reduce the bias from the specific data collected, 10 different test sets were randomly selected. The length of the trials was normalised for each phase. In total, for each combination of model type and order, 1680 fits were performed including all the test sets. The medians of the parameters resulting from individual fits on the training set were used for cross-validation.

To evaluate the quality of fit, the R-squared value of the training set and the root mean squared error (RMSE) for the test sets were evaluated. Additionally, a measure of model instability was defined. A model-order combination is considered unstable if the mean RMSE is fluctuating across the 10 test sets. In other words:

$$U = \sum_{i=0}^{9} \left| E[RMSE]_i - E[RMSE]_{i+1} \right| \tag{3.6}$$

Where $E[RMSE]_i$ and $E[RMSE]_{i+1}$ are the overall mean RMSEs resulting from the i-th and (i+1)-th test sets fitted to a given combination of model type and order, and U is the instability index: the larger U is, the less consistent is the model-order combination. The measure of model instability is used to discard those models whose performance was inconsistent due to randomness of the heuristic calculation of the parameters. The instability index is required to mitigate the effects that randomly selected test sets have on the results of fitting. The index privileges models which show a similar RMSE score on all tests sets. The assumption is that if a model truly describes a natural approaching, it is less likely that it will perform differently on different combinations of the test set. In addition, to further guide the model selection, the mean of RMSE across all the 10 test sets was evaluated. The mean RMSE is a summary of the model combination overall performance, which can be used to identify the worse performers - the model combinations with a lower overall mean RMSE. The overall variability of the

same value is also used to guide the selection if models have similar scores. This criteria is also used to loosely enforce consistency over different test sets. The criteria used to evaluate orientation model fitting were less stringent than the ones used to evaluate model fitting of the approach distance, as the motion is more complex.

To summarise, a combination of model type and order was selected based on the following criteria:

- The R-squared value of all available model combinations is compared and all models which scored less than 0.7 are discarded in both distance and orientation fits.
- The overall mean RMSE is compared across the remaining combinations. Models are discarded if they score an RMSE larger than 0.0075, when fitting distance data, or larger than 0.02, when fitting orientation data.
- The instability of each remaining combination is compared. Models are discarded if they score less than 0.0005, when fitting distance data, or less than 0.005, when fitting orientation data.
- If a clear winner does not stand out yet, worse performers are discarded.
- Variance of the overall RMSE is assessed to provide hints to guide the selection at this point.
- If two models score equally the combination with least parameters is selected.
- A simple model is also selected earlier in the process if other models with similar scores have much more parameters.

Within a selected model combination, the actual instance adopted as model for an

approach phase is the best RMSE fit across the 10 test sets. The thresholds for the R-squared, RMSE and stability values were selected in order to be strict enough to discard bad fits but not too strict to allow different aspects of a model combination to be evaluated for the best fits and to prevent over-fitting. For each approach distance phase, an approximate simplified model is also proposed. Such model is the best scoring model which does not show the same consistency over different test sets as the most reliable model does. Also, the approximate model has fewer parameters than the most reliable one. The reason why two categories of models are derived is to permit a comparison the effectiveness of different model complexities in Chapter 5. Indeed, a reliable representation of the human motion might not perform as well as a simplified one. On a robotic platform, it is more important to implement the principles of the motion rather than to imitate the human pattern. Hence a simpler model might equally reproduce the main features of human grasping despite its lower complexity, which is a desirable advantage for a practical implementation.

The model types fitted to the data were selected among the regression models available in the state of the art, by observing informally the trend of the median of the subjects' approaching data. For instance, if the data does not show oscillatory components, the preliminary fit of a sinusoidal model to the median data will be too poor to justify a thorough examination. The selected model classes for all approach distance phases and orientation quantities were Gaussian Mixtures (abbreviated to n-th order Gaussian) and Polynomial. Exponential models were tested for the third phase only, while Sum of Sines and Fourier Series were used for the fourth phase and for the orientation quantities. All the above-mentioned model classes were used to verify whether an orientation quantity is best described as a planned action. Solutions to the differential equation of first order and second order underdamped, overdamped and critically damped systems, with and without zero, were also fitted to orientation data to verify whether a quantity is best described as a reactive action.

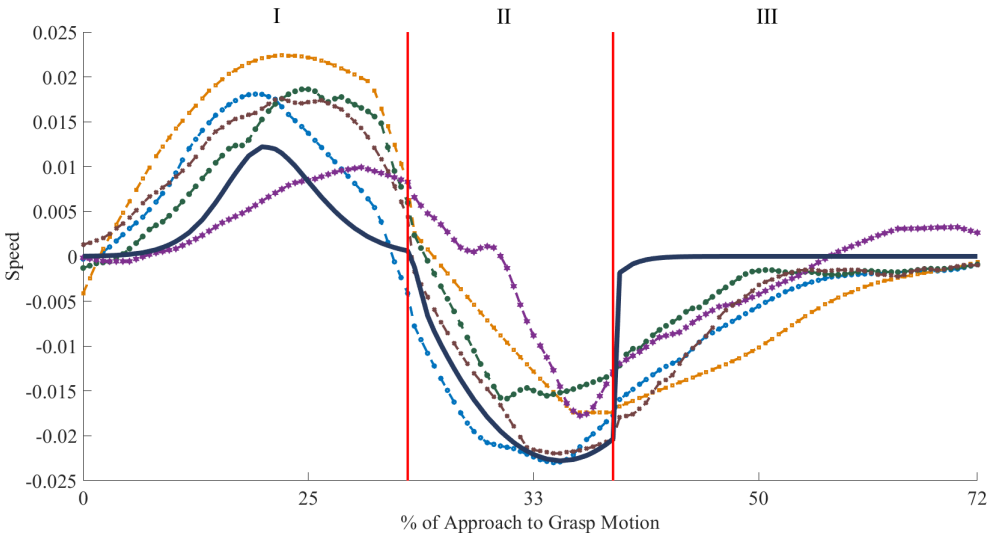


Figure 3.7: Combined output of the models fitted to each phase: Gaussian, polynomial and exponential. The model output (thick dark blue line) is overlaid on sample trials of approaching different objects performed by different subjects.

The optimal coefficients are shown in Table 3.6 for each selected most reliable and approximate model combination. Table 3.10 shows the optimal coefficients for the models of the three orientation quantities. The complete fit, resulting from the most reliable models of the first three phases, is shown in Figure 3.7.

3.4.2 Approach Distance Model Validation

First Phase

The modelling of the first phase evaluated whether the approach velocity patterns are more similar to a Gaussian, as described in [146], or to a polynomial. In addition, it was studied which complexity for each model type is required to represent most reliably the used data. In this respect, many variants of Gaussian and polynomial models were discarded. The full details of the evaluation are shown in Table 3.2. The

polynomial models were discarded as they did not fulfil the criteria set in Section 3.4.1. Those models were mostly unstable and were performing inconsistently or poorly across different test sets. This confirmed that the most reliable representation of the first phase follows a Gaussian. The best trade-off between complexity and reliability was a 3rd order Gaussian model.

The selected variant of the 3rd order Gaussian model is shown below, while the optimal coefficients are shown in Table 3.6.

$$f(t) = a_1\, e^{-\left(\frac{t-b_1}{c_1}\right)^2} + a_2\, e^{-\left(\frac{t-b_2}{c_2}\right)^2} + a_3\, e^{-\left(\frac{t-b_3}{c_3}\right)^2} \tag{3.7}$$

As alternative simplified model, the second order polynomial is the best selection. It compensates its reduced consistency with a better R-squared performance and less parameters. The optimal coefficients of the approximated model are reported in Table 3.6.

Table 3.2: Summary of model fitting results of First Phase data. Model type-order combinations with R-Squared less than 0.7 were omitted. The selected combination for the phase is highlighted in bold.

Type	Order	# Pars	R^2	E[RMSE] (10^{-3})	Var (10^{-5})	Instab. (10^{-3})
Gaussian	2	6	0.78	7.07	1.50	0.59
	3	**9**	**0.79**	**7.22**	**1.46**	**0.38**
	4	12	0.81	7.14	1.59	0.25
Polynomial	2	3	0.84	6.90	1.89	0.92
	3	4	0.93	8.75	2.50	4.67
	4	5	0.97	7.27	2.00	4.96

Second Phase

The model fitting of the second phase studied whether the human patterns are more similar to a Gaussian or a polynomial model and which complexity can appropriately describe the data. The models that meet the minimum performance criteria are listed in Table 3.3 All models showed a good RMSE value that was decreasing for some complex variants of the models. The Gaussian models were discarded due to instability and inconsistency across different test sets, poor performance or too high complexity compared to the polynomial model with similar performance. Within the polynomial models, the 2nd order polynomial was the most reliable and simple version of polynomial models, and it also showed a good R-squared performance on the training sets. Therefore this part of the motion can be represented and approximated with a polynomial:

$$f(t) = a_1\, t^2 + a_2\, t + a_3 \tag{3.8}$$

The optimal coefficients of the model are reported in Table 3.6.

Table 3.3: Summary of model fitting results of Second Phase data. Model type-order combinations with R-Squared less than 0.7 were omitted. The selected combination for the phase is highlighted in bold.

Type	Order	# Pars	R^2	E[RMSE] (10^{-3})	Var (10^{-5})	Instab. (10^{-3})
Gaussian	2	6	0.77	7.31	3.04	0.31
	3	9	0.84	7.28	3.18	0.33
	4	12	0.91	7.67	3.54	1.32
Polynomial	**2**	**3**	**0.95**	**7.25**	**3.12**	**0.06**
	3	4	0.98	7.55	2.92	1.23
	4	5	0.99	7.35	3.02	1.09

Third Phase

The nature of the motion in this phase requires a rapid convergence to near zero speed, since the hand is quickly approaching the object to finalise the grasp mostly using the fingers. For this reason, exponential models were also fitted. The form of the convergence can be used as an assessment of how likely subjects are targeting the object with a quick reactive motion stopping the hand when contact with object is achieved. The results of the fitting in Table 3.4 show that all the models admitted to the selection performed well in terms of RMSE on the test set. Further on, the Gaussian and the polynomial models were discarded since they obtained too low stability score and an inconsistent RMSE performance. Therefore the pattern of the third phase is represented by an exponential model of the second order, since the first order variant obtained an R-squared score on the edge of the minimal criteria for admission. It can be concluded that subjects do approach the object with a quick and direct reactive motion rather than with a planned motion as for the other phases.

The final Exponential model is shown below, while the optimal coefficients are shown in Table 3.6.

$$f(t) = a_1\, e^{b_1\, t} \; + \; a_2\, e^{b_2\, t} \tag{3.9}$$

The structure of the motion is comparable with the step response of a second order over-damped spring-mass-damper system, as the exponents of both terms are negative and less than 1 as per definition. However, the steady state gain is not equal for both terms but it differs by a factor of 10 for each exponential. This suggests that the settling dynamics is similar to a second order system but the steady state differs. However, in this case, once the data reaches the steady state, the execution of the fourth phase

begins.

The third phase can also be approximated with a simplified model. A second order polynomial compensates for its inconsistent performance across test sets with an improved RMSE mean score and a loss of a parameter. It is important to stress that a second order polynomial is a preferred option for an approximate model at this point, as the other phases can also be approximated with a second order polynomial. Therefore, by combining the three polynomials of the same order together it is also easier to switch between the models during the evolution of the motion on the robot. The optimal coefficients of the approximated model are reported in Table 3.6.

Table 3.4: Summary of model fitting results of Third Phase data. Model type-order combinations with R-Squared less than 0.7 were omitted. The selected combination for the phase is highlighted in bold.

Type	Order	# Pars	R^2	E[RMSE] (10^{-3})	Var (10^{-5})	Instab. (10^{-3})
Gaussian	3	9	0.69	4.49	1.28	0.46
	4	12	0.71	4.46	1.14	0.55
Polynomial	1	2	0.69	4.41	1.19	0.28
	2	3	0.85	4.29	1.23	0.65
	3	4	0.93	4.70	1.04	4.41
	4	5	0.97	5.01	1.21	6.56
Exponential	1	2	0.69	7.40	1.03	0.04
	2	**4**	**0.81**	**7.41**	**1.03**	**0.13**

Fourth Phase

It has been already discussed in Section 3.3.2 that this phase is not common across different subjects, however it is discussed here for completeness. The phase has been analysed to describe the structure of this part of the motion. In many instances, the data shows oscillatory components, hence Fourier series and Sums of Sines were fitted to the data to capture this feature. Such models are not appropriate for the other phases

as they do not exhibit an oscillatory behaviour. As shown in Table 3.5, all the Gaussian models, the 1st and 2nd order polynomial and the Sums of Sines, have an R-squared performance too low to be considered. This suggests that the hand undertakes complex oscillatory components rather than a period of relative immobility. Of the remaining models, a 1st order Fourier series is selected as the most reliable representation of the motion. The other models have a bad RMSE performance or are too complex, such as Fourier series. The model has the following structure and its parameters are listed in Table 3.6:

$$f(t) = a_1 + a_2 \cos(c_2\, t) + b_2 \sin(c_2\, t) \tag{3.10}$$

The polynomial models show low stability as their performance is inconsistent across different test sets. To provide further information, a third order polynomial can capture the prominent characteristics of this phase as an approximate representation. This model compensates for its inconsistency with an improved RMSE performance and its optimal coefficients are reported in Table 3.6. It is also possible that the observed oscillations are the result of a closed loop control performed to complete the grasping, which could explain also the inter subject variability. However this type of analysis is beyond the scope of this chapter and can be performed in future works.

Table 3.5: Summary of model fitting results of Fourth Phase data. Model type-order combinations with R-Squared less than 0.7 were omitted. The selected combination for the phase is highlighted in bold.

Type	Order	# Pars	R^2	E[RMSE] (10^{-3})	Var (10^{-6})	Instab. (10^{-3})
Fourier	**1**	**4**	**0.75**	**3.71**	**6.26**	**0.88**
	2	6	0.90	3.91	6.07	1.04
	3	8	0.95	3.71	5.97	0.80
	4	10	0.98	3.92	5.84	1.97
Polynomial	3	2	0.78	3.38	6.65	1.24
	4	4	0.85	3.40	6.45	9.00

Table 3.6: Optimal coefficients of the most reliable and approximated models selected for each phase

Coefficients	**Phases**			
	First	Second	Third	Fourth
	Most Reliable Models			
a_1	$1.62\,10^{-3}$	$0.41\,10^{-4}$	$-1.73\,10^{-4}$	$1.13\,10^{-3}$
b_1	26.18	N/A	-0.7063	N/A
c_1	4.95	N/A	N/A	N/A
a_2	$5.77\,10^{-3}$	$-18.08\,10^{-4}$	$-30.89\,10^{-4}$	$-9.22\,10^{-4}$
b_2	26.97	N/A	-0.4122	$2.09\,10^{-4}$
c_2	8.396	N/A	N/A	1.51
a_3	$5.246\,10^{-3}$	$-30.16\,10^{-4}$	N/A	N/A
b_3	29.65	N/A	N/A	N/A
c_3	12.05	N/A	N/A	N/A
	Approximated Models			
a_1	$-0.12\,10^{-4}$	$0.41\,10^{-4}$	$-5.18\,10^{-6}$	$2.98\,10^{-7}$
a_2	$5.93\,10^{-4}$	$-18.08\,10^{-4}$	$5.59\,10^{-4}$	$-1.18\,10^{-5}$
a_3	$-3.92\,10^{-4}$	$-30.16\,10^{-4}$	$16.98\,10^{-4}$	$1.31\,10^{-4}$
a_4	N/A	N/A	N/A	$-4.40\,10^{-4}$

3.4.3 Approach Orientation Models Validation

Azimuth of Approach

As shown in Table 3.7, the solutions to the differential equations of first and second order systems do not produce good fits of the Azimuth angle of approach with a R-squared score sufficient to be considered as a model for this quantity. Therefore, it can be concluded, that the motion of the Azimuth of approach is planned and intentional action. However, all the candidate models for the Azimuth describe an oscillatory motion. Fourier series is the only class of models whose consistency and RMSE values are within the set thresholds over different test sets. The simplest best model is a Fourier series with 4 elements. The model structure is shown below and the parameters as listed in Table 3.10:

$$f(t) = a_1 \cos(t\,f) + b_1 \sin(t\,f) + a_2 \cos(3\,t\,f) + b_2 \sin(2\,t\,f) + b_3 \sin(4\,t\,f) \quad (3.11)$$

Table 3.7: Summary of model fitting results of Azimuth Angle of Approach data. Model type-order combinations with R-Squared less than 0.7 were omitted. The selected combination for the phase is highlighted in bold.

Type	Order	# Pars	R^2	E[RMSE] (10^{-2})	Var (10^{-5})	Instab. (10^{-3})
Fourier	**4**	**10**	**0.81**	**1.75**	**5.60**	**1.31**
	5	12	0.89	1.76	5.54	1.43
	6	14	0.94	1.76	5.78	2.35
Sum of Sine	4	12	0.73	2.73	2.69	4.83
	5	15	0.77	2.61	2.87	6.94
	6	18	0.77	2.54	2.99	5.73

Hand Rotation

The results of fitting for the hand rotation data show that the best model class for the planned action hypothesis are Fourier series of even order. The Polynomials score a high error, while Sums of Sine have higher model complexity than the Fourier series as shown in Table 3.8. The fits to the step response of an underdamped Second Order System with and without zero also scores well in terms of stability of the performance and overall error. Therefore, it is possible to compare the two hypotheses for the hand rotation. Both Second Order Systems outperformed the Fourier series models in terms of consistency over different test sets. The only Fourier series with a comparable RMSE score has 14 parameters, which is overcomplicated with respect to the two Systems. Therefore, it is possible to conclude that the hand rotation is a reactive action. An underdamped Second Order System without a zero best describes the motion pattern, since it has the same RMSE and consistency values of the other System but a parameter less.

The model has the following structure:

$$f(t) = a_1 \left(1 - a_2 \, e^{-b_2 \, t} \, \sin(f \, t + c_2)\right)) \tag{3.12}$$

Where a_1 is the static gain, while a_2, b_2, c_2 and f are defined as follows:

$$a_2 = \frac{1}{\sqrt{1-\zeta^2}} \tag{3.13}$$

$$b_2 = -\zeta\omega_n \tag{3.14}$$

$$c_2 = 1-\zeta^2 \tag{3.15}$$

$$f = \arccos(\zeta) \tag{3.16}$$

The parameters of the system are the natural frequency (ω_n) and the damping ratio (ζ). Table 3.10 reports the parameters of the system as an aggregate to follow the structure expressed in Equation 3.12.

Table 3.8: Summary of model fitting results of Hand Rotation data. Model type-order combinations with R-Squared less than 0.7 were omitted. The selected combination for the phase is highlighted in bold.

Type	Order	# Pars	R^2	E[RMSE] (10^{-2})	Var (10^{-5})	Instab. (10^{-3})
Polynomial	5	6	0.74	$6.9\,10^6$	1.16	$3.6\,10^9$
	6	7	0.81	$3.6\,10^8$	9.65	$6.9\,10^7$
Fourier	3	8	0.80	1.25	0.78	7.06
	4	10	0.88	1.17	0.84	3.69
	5	12	0.93	1.10	0.89	11.82
	6	14	0.97	0.99	1.09	2.87
Sum of Sine	3	9	0.73	1.75	0.55	8.38
	4	12	0.79	1.54	0.60	3.85
	5	15	0.72	1.51	0.60	2.86
	6	18	0.81	1.39	0.66	2.42
2nd Order	**2**	**2**	**0.81**	**0.95**	**1.40**	**0.01**
2nd Order w/t Zero	2	3	0.79	0.95	1.40	0.02

Zenith of Approach

The results of the fit of the Zenith angle of approach disproved the hypothesis that this quantity is controlled as a reactive motion, as it is shown in Table 3.9. Indeed, no solution to the differential equation of first or second order models produced a fit with an R-squared above the set threshold. It can be concluded that this motion is a planned and intentional action. Across the admitted model classes, Gaussian models describe it best, since such models have a performance similar to Fourier Series and a higher stability across different test sets. Since all the Gaussian models perform consistently across the test set, and their stability score near zero, the simplest valid version of the class was selected. This corresponds to a 3rd order Gaussian model with parameters shown in Table 3.10, its structure is shown below:

$$f(t) = a_1\, e^{-((t-b_1)/c_1)^2} + a_2\, e^{-((t-b_2)/c_2)^2} + a_3\, e^{-((t-b_3)/c_3)^2} \tag{3.17}$$

Table 3.9: Summary of model fitting results of Zenith Angle of Approach data. Model type-order combinations with R-Squared less than 0.7 were omitted. The selected combination for the phase is highlighted in bold.

Type	Order	# Pars	R^2	E[RMSE] (10^{-2})	Var (10^{-5})	Instab. (10^{-3})
Gaussian	**3**	**9**	**0.71**	**1.35**	**9.30**	$1.15\,10^{-8}$
	4	12	0.74	1.36	9.29	$1.50\,10^{-13}$
	5	15	0.77	1.35	9.30	$7.06\,10^{-11}$
	6	18	0.76	1.36	9.29	$2.52\,10^{-13}$
Fourier	3	8	0.78	1.37	8.78	1.67
	4	10	0.88	1.38	8.68	1.87
	5	12	0.94	1.34	8.99	0.71
	6	14	0.96	1.34	9.12	0.60
Sum of Sine	4	12	0.73	1.72	7.44	7.58
	5	15	0.77	1.64	7.69	5.19
	6	18	0.77	1.65	7.75	7.44

Table 3.10: Optimal coefficients of the models selected for orientation quantities. The coefficient of the Second Order System are the result of the multiplications of the ω_n and ζ parameters of the two poles.

Coefficients	**Quantity**		
	Azimuth	Rotation	Zenith
a_1	7.60 10−4	-3.34 10−05	0.03
b_1	2.14 10−4	N/A	-0.83
c_1	N/A	N/A	0.22
a_2	-3.41 10−4	3.38	0.02
b_2	2.84 10−4	1.61	-0.55
c_2	N/A	0.79	0.26
a_3	N/A	N/A	0.01
b_3	2.33 10−4	N/A	-0.48
c_3	N/A	N/A	0.29
f	1.40	1.89	N/A

3.5 Discussion

It is commonly agreed that the approaching to grasp motion follows a pre-defined timed plan, in terms of hand transportation motion and grip formation, which can be perturbed within limits [147, 148].

It is worth observing that the first phase of the analysed data has a bell-shaped form. This result is in line with many findings such as [146] and [149]. Specifically, authors in [146] also fitted a Gaussian model to their data as it is done in this study. However the complexity of the model was higher due to the fact that the whole approaching motion was involved. The study in [146] suggests that reaching is an open-loop motion. Our findings, however, demonstrate that this open-loop profile terminates before the end of the motion. Marteniuk et al. [150] also observed a similar difference when subjects were asked to reach to a point or approach to grasp for lifting the object. The authors found that the deceleration of the hand is longer for more complex tasks. Our findings also confirm the difference between reaching to a point and approaching to grasp.

Indeed, the open-loop reaching part of the motion has a defined duration, after which the strategy of the approach motion is defined. In this regards, the second phase is the moment when the final approach and grasping patterns are finalised. In agreement with Marteniuk et al. [150], this phase features a sharp deceleration with a common shape across subjects, possibly because the subjects performed the same task. It was found that the actual length of the phase is different for each subject and object grasped. This might be caused by contingent factors during the experiment. The results also show that most of the finger preshaping motion is performed in this phase. This finding is in line with observations in [151], and it complements the author's previous work [8] that describes how fingers are displacing for grasping. It is possible to observe that the precise approaching to grasp strategy is determined by the end of this phase and performed in the next.

Jeannerod [151], in a similar study involving approach to grasp for transporting, also observed that subjects undergo a low-velocity phase consistently at the same moment near the end of the motion. The study suggests that this phase is necessary for prehension and is not a corrective action. The results presented in this Chapter add to this statement. The presented data demonstrates that the last phase of approaching to grasp is finalising the grasping strategy determined while approaching. The findings also demonstrate that the third phase is the only reactive part of the motion where the variability of finger joint speed is minimal. This suggests that the termination of the approach to grasp motion is a scripted mechanism.

Concerning the orientation, it is possible to observe that both approach angles, Azimuth and Zenith from the object's axis of rotation, are part of a planned approach strategy. Models describing a reactive action performed poorly and could not be considered for comparative evaluation. Specifically, azimuth of approach has clear oscillatory components and, as such, requires a complex model. The hand rotation, instead, is clearly

better described as a reactive action with dynamics similar to an underdamped mass-spring-damper system. Human muscles are also modelled as underdamped systems in the state of the art [152], which supports this finding. Since the hand rotation does not describe an object-hand relationship as the other angles of approach, it is logical to conclude that the data mostly captures the activity of the arm muscles.

The presented findings support the hypothesis that the approach to grasp motion follows mostly a planned strategy, although the last phase of the motion and the hand rotation are scripted and reactive components. Additionally, the data supports the statement that the approaching plan is defined at a precise moment during the course of the action, rather than from the beginning of it. The results support the hypothesis that the finger motion is synchronised with the hand motion, as most of the preshaping is performed in one specific phase.

3.6 Conclusions

In this chapter, two components of a grasp affordance were defined: an initial approaching to grasp phase, and a second phase where the desired grasp pattern is implemented. The approach to grasp for hammering was studied, collecting data from 9 subjects who used very different objects, in different orientations, to perform hammering action. The collected data was analysed and precise and approximate mathematical models were selected to describe the motion.

The main contributions of this chapter are summarised below:

1. Subjects share a common approaching to grasp pattern that can be divided in approach distance and three orientation quantities: Hand Rotation and Azimuth and Zenith Angle of Approach to the object.

2. The approach distance has a clear structure with four phases. The Rotation and Zenith patterns of approach are object and trial independent, but subject dependent. The structure of the Azimuth patterns of approach is, however, subject independent as well.

3. The first two phases of the pattern of approach to grasp can be reliably modelled as a planned and intentional motion. The third phase follows the dynamics of a spring-mass-damper system and a reactive motion pattern.

4. As alternative hypothesis, the entire approaching motion can be approximated to a fully planned motion modelled as a second order polynomial, to simplify a technical implementation.

5. The fourth phase of the action, where the hand is adjusted to implement the grasping, presents oscillatory components and a pattern that is not common across different subjects.

6. The Azimuth and Zenith angles of approach can be modelled as a planned motion, while the hand rotation is best represented as a reactive motion following the dynamics of a second order underdamped system.

Chapter 4

The Role and Kinematics of the Thumb in Human and Robotic Grasping

4.1 Introduction

A grasp affordance action is composed of an approaching phase that appropriately positions the palm near the object, and a grasping phase where fingers envelop and restrain the object in the hand. Those two phases complement each other, as poor positioning of the palm can be compensated by smart restraining of an object, while a simplistic grasping algorithm can perform well if the palm is aligned appropriately. An ideal robotic hand, as dexterous as a human one, would allow less strict requirements on the alignment of the palm. Indeed, its fingers would be able to passively restrain the object in hand as easily as humans do. Since robotic fingers are neither as compliant nor as dexterous as the human ones, it is harder to safely restrain an object in the

hand if the palm is not aligned correctly. Therefore, it is required to quantify the kinematic mismatch between human and robotic hands in order to define how much precision is required in the approaching phase. This is done to align the palm so that the grasping phase can succeed. Additionally, a quantification would highlight which grasp affordances are too difficult to implement due to the current kinematic limitations of robotic hands.

This chapter examines the motion of human fingers in grasping to understand whether the motion of the thumb influences the shaping of the grip independently from the specific object being grasped. Additionally, the reachability space of state-of-the-art thumbs of robotic hands is analysed and related to the human reachability space to understand whether the current design limits the available finger posture and the set of grasp affordances that can be easily replicated in robotics. Suggestions for improving the thumbs of robotic hands are provided.

The motion of five fingers of the human hand is analysed in this chapter. However, the study of the role of the thumb in grasping is in the focus of this study.

This choice of studying the role of the thumb in grasping is motivated by evolutionary and anatomical reasons. It is well known in the community of evolutionary biology [15,153] that the opposition of the thumb towards the other fingers is the most important advancement that granted manipulation skills to prehistoric humans and primates. The ability of thumb opposition is an infrequent skill in nature, mostly developed in humans and some primates in different ways [18,19]. The above-mentioned studies highlight the importance of the opposition feature of the thumb but do not quantify the characteristics of its motion.

Anatomical studies of the hand and the thumb were conducted to understand the joint structure and the mechanisms of actuation of the fingers [154]. For instance, the oppo-

sition mechanism of the thumb is assessed through anatomical analysis of hand skeletal bones [35], or analysis of the variability of the trapeziometacarpal motion [36]. It is well acknowledged that the loss of the thumb corresponds to the loss of 40% of the hand function [16]. However, a common agreement on the mechanics of the human thumb opposition was not reached yet [37].

In robotics, several attempts to model the kinematics of the human hand and the thumb were done in the past, settling the number of joints of the hand from 15 [73] to 25 [74]. However, there is no general agreement on the best solution to create a model that renders the dynamics of human grasping and manipulation. Consequently, very different approaches are used to design robotic hands. In addition, biological or technical explanation is rarely given for different design choices. In other studies, the grasping capabilities of robotic grippers and human hands are evaluated and compared in order to understand better the main principles of grasping. For instance, [115] is addressing solutions to the correspondence problem for the entire hand. Such contributions are very important, but they don't take into account the prominent role of the thumb.

Based on the existing literature, this chapter addresses the following questions:

1. Is the displacement of the opposing thumb in humans a determining factor for shaping the grip independently of the grasped object?

2. How far does the reachability space of state-of-the-art robotic thumb kinematics span evenly on the human thumb reachability space?

3. How is it possible to improve the design of state-of-the-art robotic thumbs to enhance the hand dexterity?

This chapter is organized as follows: Section 4.2 presents the methodology of the analysis and proposes a classification of grasping postures. In Section 4.3, the first question on

the role of the thumb in grasping is investigated by analysing the fingertip motions. Then, in Section 4.4 the reachability space of the human thumb is compared with the reachability spaces of two robotic thumbs, and the representation of each grasp category is evaluated. Section 4.5 discusses the results of the comparison and provides some suggestions for improving the thumb kinematic design. Finally, Section 4.6 concludes the chapter and summarises the results.

4.2 Methodology

4.2.1 Methodology of the Analysis

To explore the role of the thumb in grasping, it is required to analyse the position of all fingertips for different postures across subjects. The dataset used for the analysis is described in the following subsection. The position of the fingertips is expressed in the coordinate system of the defined reference frame. The centre of the reference frame is placed on the dorsum of each subject's hand, so that any motions of the wrist and the arm are eliminated. Figure 4.1 shows the orientation of the axes of the reference frame. The position of the fingertips is analysed and compared in three-dimensional space. Therefore, the rotation and translation components of each data point are combined together in relation to the reference frame. For the purposes of analysis it was decided to use the displacements of the fingertips only, without the additional information that a kinematic model of the hand can give. The use of a kinematic model imposes an arbitrary structure to the data that biases analysis reducing its generality. Such decision is also motivated by the fact that there is no agreement on the ideal kinematic model of the hand as mentioned in Section 4.1.

The statistical analysis performed for these studies is based on the evaluation of the

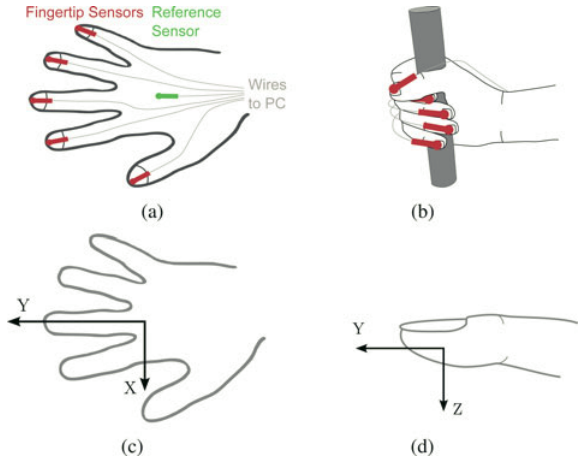


Figure 4.1: Description of the set up of the sensors [115]. (a) Position of the sensors on the fingers (red) and of the reference sensor on the dorsum (green). (b) Example of grasp posture. (c) and (d) Coordinate system used in the analysis.

standard deviation, covariance and mode of covariance of the displacements for all digits across subjects. The full demonstration sequence — from approaching the object to retreating the hand — is taken into account to analyse the complete motion of the fingertips during a grasping action. The analysis of variance (ANOVA) is used to evaluate whether there are factors that can significantly influence the displacement of the fingers, namely, the grasped object, the chosen posture, and the individual behaviour of each subject. The statistical evaluation was performed across 300 trials. Before this analysis, a Kolmogorov-Smirnov test was performed to test whether the data follow a Gaussian distribution. The results shown that the data is normally distributed for all trials. In the ANOVA tests, the impact of the factor was considered significant in case the null hypothesis was rejected with 95% confidence level that corresponds to probability distribution (p) less than 0.05. One-way ANOVA tests were performed, with one inter-cells degree of freedom and 150 within-cell degrees of freedom ($F_{(1,150)}$).

4.2.2 Description of the Dataset

To perform this study it is required to analyse as many different grasp postures as possible. In this way, a more general rule that applies to a large variety of grasping postures can be formulated. For this reason, data available in the GRASP database [38] was used for the analysis. The purpose of this dataset is to create a comprehensive taxonomy of grasp postures. This chapter, instead, defines the role of the thumb in shaping human hand grasps and compares the freedom of movement of the thumb with the reachability space of the robotic thumbs.

The GRASP database captures 31 different grasp postures, and, therefore, the fingers span the majority of the possible positions used in grasping. As described in [115], five different subjects, all right handed, were asked to perform the grasp postures according to the classification outlined in [155]. The subjects, three males and two females, have an average hand length of 185.2 mm and hand width of 81.1 mm with standard deviation of 13.3 mm and 7.4 mm respectively. The objects used to perform each posture are listed in Table A.1 in the Appendix A. The subjects positioned the hand flat open with the dorsum up in front of the table. They were requested to replicate a grasp configuration as shown in a picture or, in case of difficulties, to repeat it from a demonstration. During the grasping experiment, the hand approached the object, lifted it and retreated back to the initial position after the object was placed down.

Each demonstration was recorded using Polhemus Liberty system with six magnetic sensors. The spatial and angular resolution of each sensor was 0.8 mm and 0.15 degrees respectively. A sensor was applied on each fingernail and, in addition, one sensor was placed on the dorsum of the hand, acting as reference point for the captured data as shown in Figure 4.1. The movements of the hand were recorded at 240 Hz. Each subject grasped 14 objects twice, each trial had 600 uniformly sampled data points.

4.2.3 Thumb-oriented Classification

The GRASP database uses a list of 31 different classes that are partially derived from Cutlosky's posture classification [22]. In this chapter, it was decided to simplify the list of possible postures to four, based on the employment of the thumb. This approach is required to highlight the role of the thumb in grasping for this analysis, and it was adopted from the classifications outlined by Pouydebat et al. [18] and by Napier [39]. Originally, Napier divided humanoid grasping in two categories - power and precision grasps. Pouydebat et al. extended this classification to five categories based on the contact surface. This classification takes into consideration the placement of the thumb with respect to the other four fingers. It also takes into account whether the grasp is performed using the fingertips or the whole surface of the digits. The proposed classification uses thirty grasp postures, and each posture fits to one specific class out of four based on the functionality of the thumb.

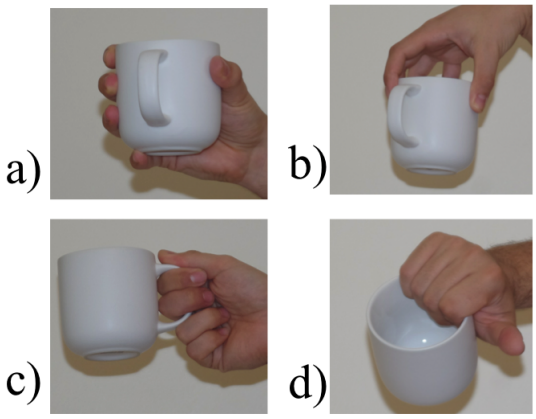


Figure 4.2: Four posture classes based on the position of the thumb, a) Power Grasp, b) Precision Grasp, c) Key Grasp, d) Primate Grasp

A comprehensive list of examples of the described classification schema, and a comparison with the GRASP classification, is shown in Table A.1 of Appendix A. The formulated posture classes are shown in Figure 4.2 and are defined as follows:

1. **Power Grasp:** the thumb opposes the other four fingers; the object is touched with the whole surface of the digits.

2. **Precision Grasp:** the thumb opposes at least one of the other four fingers; the object is touched with the fingertips only.

3. **Key Grasp:** the thumb does not oppose any other fingers but it is still used for prehension (grasping). There are no additional assumptions in respect to the other fingers.

4. **Non-human Primate Grasp:** the thumb is not used for grasping - its fingertip and most of the surface of this finger is not in contact with the object; the object is grasped using any of the other four fingers in any configuration. This posture class, along with the Power Grasp, is very popular among non-human primates but it is less frequently used by humans.

4.3 Analysis of Thumb Motion

4.3.1 Analysis of Variability of the Movements of Fingers

As a first step, the variability of each finger across postures and grasped objects was analysed in order to explore the role of the thumb in prehensile grasping. The standard deviation was calculated as follows. Initially, the 3D position of a finger was evaluated:

$$P_{f_{3\times1}} = R_{f_{3\times3}} T_{f_{3\times1}} \tag{4.1}$$

Where P_f is the 3D position of one data point of finger. f, T_f and R_f are the translation and rotation components of finger f expressed in the reference frame of the dorsum. The

calculation was performed on all data points for all trials of the dataset. Afterwards the norm of each 3D position was calculated:

$$N_f = \|P_{f_{3\times1}}\| \tag{4.2}$$

Where N_f is the Euclidean norm corresponding to 3D position P_f. The calculation was performed on each 3D position of the fingers for all the trials of the dataset. Finally, the trials were divided by grasping class, as indicated in Table A.1, and the standard deviation was calculated on the trials of a grasping class for each finger individually. Figure 4.3 shows the standard deviation of the norms of the three-dimensional displacement of the fingers across the whole dataset. The use of the norm allows taking into account both the magnitude and orientation of the positions of the fingertip with respect to the origin. Every coloured bar represents the variability of movements for each finger in a given grasp class. It is considered that the thumb is the most mobile digit of the hand [154]. However, Figure 4.3 demonstrates that the magnitude of the variability of the motion of the thumb is the lowest across the four postures. The other fingers have more overall variability of motion: the index finger - 95% of variability more than the thumb, the middle - 78%, the ring - 64%, and the little - 50%.

To evaluate the displacement of the thumb in relation to the other digits further, as well as to observe the influence of a posture class, the covariance value of fingertip-to-fingertip positions was calculated for each trial of the dataset. The covariance for a single trial was calculated as follows:

$$C_{f_i,f_j} = E[(N_{f_i} - \mu_{N_{f_i}})\,(N_{f_j} - \mu_{N_{f_j}})] \tag{4.3}$$

Where N_{f_i} and N_{f_j} are the norms of the fingertip displacements of digit f_i and finger

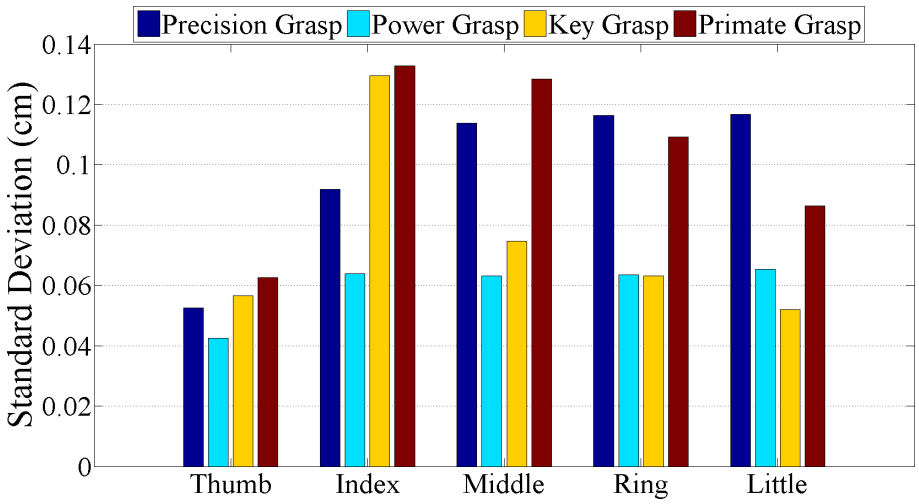


Figure 4.3: Standard deviation of finger movements across the entire GRASP dataset [115]. The contribution given by each grasp posture class to its overall standard deviation is highlighted in different colour for each finger.

f_j respectively for a single trial. $\mu_{N_{f_i}}$ and $\mu_{N_{f_j}}$ are the means of finger f_i and finger f_j respectively for the trial. C_{f_i,f_j} is the covariance between digit f_i and digit f_j. The covariance values were calculated among all the five fingers. For each finger, the covariance values of the trials were grouped by grasping class, as indicated in Table A.1 and summed up to improve the representation:

$$S_g = \sum_{k=1}^{n} C^k_{f_i,f_j} \tag{4.4}$$

Where S_g is the summed covariance for grasping class g. $C^k_{f_i,f_j}$ is the covariance between finger f_i and finger f_j for trial k. n is the number of trials belonging to grasping class g. The results are shown as a histogram in Figure 4.4. Each sub-figure shows the results for the corresponding posture class, and each bar of different colour represents a different finger. A lognormal distribution was fitted to the data to simplify the visual comparison of the covariance values across fingers and posture classes. A lognormal distribution was chosen since it produced the best fit and the data is normally distributed. As a result,

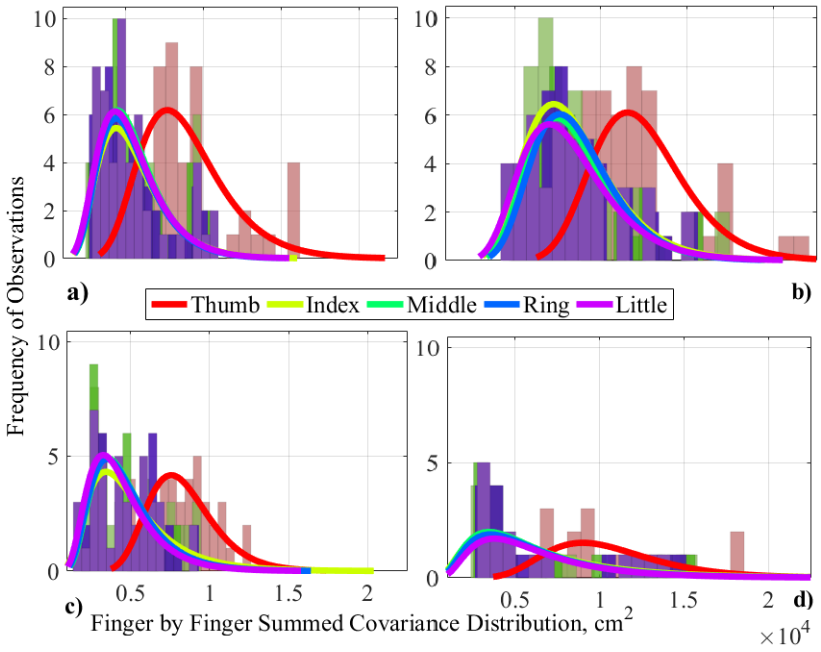


Figure 4.4: Summed covariance between one fingertip and the fingertips of the other digits is displayed on the histograms along with the fitted distribution curve across all subjects and grasped objects. Figure *a)*, *b)*, *c)* and *d)* represents the results for Power, Precision, Key and Primate Grasp classes respectively. Each colour represents a different digit. The closer the bars and the curves are to zero, the less covariated is a specific finger while performing a specific grasp type. The higher a bar is, the more frequent is that specific motion coupling between a finger and the other digits. Each curve is a lognormal distribution fitted on the data of the GRASP dataset [115].

Figure 4.4 highlights the strong magnitude of covariance between the thumb and the other digits. In other words, the motion of the thumb is leading the displacement of the other fingers. In addition, this trend can be observed in all of the four posture classes. The thumb is the digit that shows less variability of motion across the four posture classes. Therefore, it can be concluded that the contribution to high covariance values of the thumb motion comes primarily from the other four fingers. This suggests that the motion of the other four digits follows the motion of the thumb.

To further validate the leading role of the thumb for each grasping posture, the most frequent covariance values (modes) per finger and grasping posture, obtained from the

histogram (Figure 4.4), are compared. The mode is the value occurring most frequently in the distribution. The mode of covariance distribution for each finger was calculated and shown in Figure 4.5. The diagram confirms that the motion of the thumb is more interdependent compared to the motion of the other fingers.

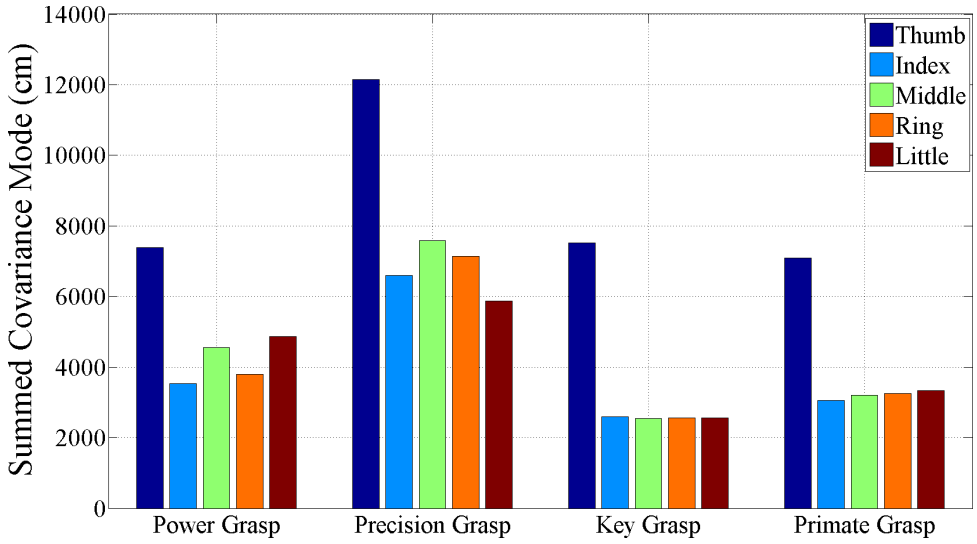


Figure 4.5: Most frequent covariance (mode) per finger and posture class. The mode is calculated on the summation of the covariance for each posture class.

In summary, the analysis in this section shows that the movement of the thumb during grasping is less when compared to the displacement of the other fingers (Figure 4.3), but at the same time the displacement of the other fingers strongly depends on the displacement of the thumb (Figure 4.4). In addition, the motion of the thumb is the most interdependent with the other fingers (Figure 4.5). These findings demonstrate that the displacement of the thumb is a determining factor for the motion of the other fingers in grasping.

4.3.2 Statistical Evaluation

To verify the results on the prominent role of the thumb statistically, ANOVA tests were performed. The statistical analysis was performed to verify whether additional conditions influence the variance of motion of each finger. In this case, the additional conditions are - the object being grasped, the subject performing the grasp, and one of the four assigned postures. The purpose of this analysis is to add further evidence to the findings in Section 4.3.1 that state that the thumb motion has a prominent role in shaping the hand grip. Results of the ANOVA test are shown in Table 4.1.

Table 4.1: Summary of ANOVA test results on the GRASP database [115]. F is the value of the Fisher's index, while p is the P-value. Entries in bold are statistically significant.

Finger	Posture		Subject		Object	
	F	p	F	p	F	p
Thumb	**5.94**	**<0.02**	0.28	0.6	0.61	0.44
Index	**3.58**	**<0.01**	**3.05**	**<0.01**	**7.91**	**<0.0001**
Middle	1.62	0.2	0.29	0.59	**5.46**	**<0.001**
Ring	**3.48**	**<0.01**	3.05	0.92	0.75	0.39
Little	**4.35**	**<0.01**	0.65	0.42	0.07	0.79

The strongest dependency on the variability of the thumb position is the posture executed by a subject to grasp an object ($F_{1,150} = 5.94,\ p < 0.02$). The object ($F_{1,150} = 0.61,\ p = 0.44$) and the subject ($F_{1,150} = 0.28,\ p = 0.6$) are statistically not influencing the thumb position. The displacement of other fingers, such as the index finger ($F_{1,150} = 3.58,\ p < 0.01$), the ring finger ($F_{1,150} = 3.48,\ p < 0.01$) and the little finger ($F_{1,150} = 4.35,\ p < 0.01$), are also playing a significant role. In addition, it can be observed that there is a strong dependency on the object for index ($F_{1,150} = 7.91,\ p < 0.0001$) and middle ($F_{1,150} = 5.46,\ p < 0.001$) fingers. Such dependency on the object might be used to interpret the specific geometry of the grasped object.

Similarly to the thumb, the ring and little fingers depend only on a specific posture. However, these two fingers are only employed in 63.3% and 53.3% out of all the grasping demonstrations respectively. Conversely, the thumb is used in 90% of the grasping executions. Therefore, the results of the statistical evaluation highlight the broad use of the thumb that is independent from the grasped object. In other words, similar motion of the thumb can be used to grasp different shapes.

To summarize, the index and the middle fingers can provide information on the object geometry, while the thumb motion can determine which grasp posture is used. The posture can be used to grasp a possible set of objects. For instance, the same thumb motion is suitable to grasp to transport a cylindrical pole and a flat rectangular crowbar. The index and middle fingers can distinguish whether the subject is grasping the crowbar or the pole.

4.4 Human-Robot Thumb Reachability Space Comparison

4.4.1 Motivation

The goal of this section, it is to investigate whether state-of-the-art robotic thumbs are representing the motion of the human thumb evenly for each posture class. It is required to evaluate the reachability space of robotic thumbs compared to the human thumb. The iCub [62] and the Shadow [70] robotic hands were used for the analysis. The main reason behind this choice is that those two hands model the kinematic of the thumb similarly to other multifingered human-like robotic hands. For instance, the Robonaut 2 Hand [156], Gifu Hand [83], HIT/DLR Hand 2 [63] and Sandia Hand [89] are some of the examples. The kinematic chain of the thumb of the above mentioned hands is inspired by the morphology of the human thumb. The human thumb can

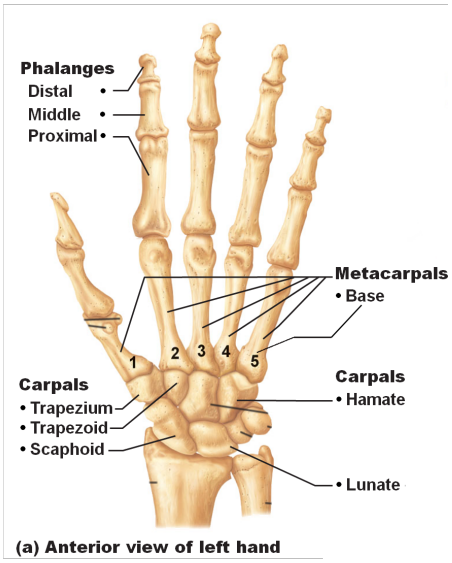
oppose to the other fingers, as the region of the trapezium, trapezoid and scaphoid can be folded [154]. It can also be adducted and abducted around the distal and proximal articulations of the fingers. Adduction, abduction and opposition were the motions analysed. Additionally, the two selected hands are either commercially available or publicly accessible as scientific prototype; in addition, their kinematic design is freely available for analysis.

This chapter considers robotic hands with a thumb opposition joint placed at the base of the thumb proximal phalanx, near the articulation of the thumb metacarpal (see Figure 4.6(a)). The same kinematic design of the opposition joint of the thumb is used in the above-mentioned hands. The iCub hand has a four Degrees of Freedom (DoFs) thumb, as shown in Figure 4.6(b), and the Shadow hand has a five DoFs thumb, as shown in Figure 4.6(c). The analysis of the reachability space of these two robotic thumbs can give a better understanding of the rendering of the thumb motion in robotics.

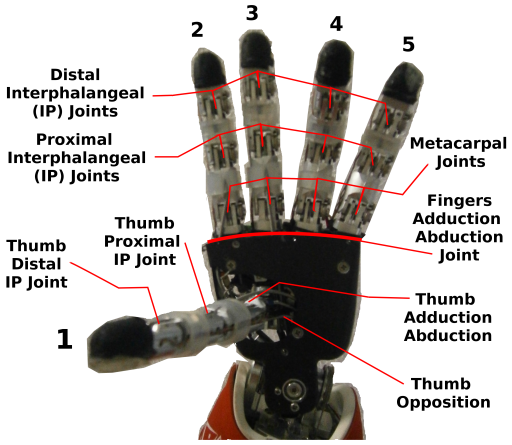
It is worth mentioning that other human-inspired multifingered grippers, such as the Barrett Hand [69] or the metamorphic hand [157] are not considered in this chapter. This is due to the use of a grasping principle different from human morphology. For instance, the fingers of the Barrett Hand synchronously rotate around the palm surface, and the metamorphic hand has a reconfigurable metamorphic palm that allows complex configuration of the fingers different from human morphology.

4.4.2 Methodology of Analysis

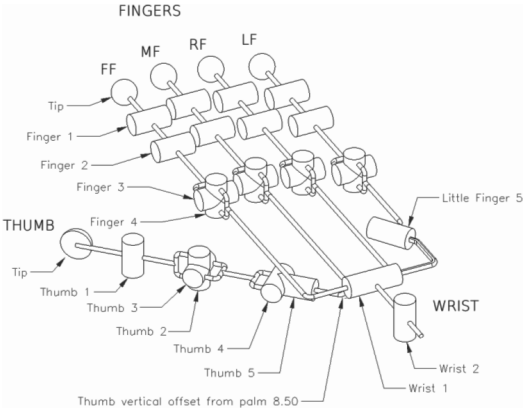
In this part, the methodology used to compare the reachability spaces of human and robot thumbs is described. The first step for the comparison of reachability spaces of the human and robot thumbs is to calculate all the positions of the tip of the robotic thumb in 3D space. The forward kinematics of the robotic thumbs were calculated numerically



(a) Schema depicting the 26 bones of the human hand. Bones referenced in text are reported on the figure.



(b) Diagram of the iCub hand kinematics.



(c) Diagram of the Shadow hand kinematics.

Figure 4.6: Diagrams of human hand skeleton and the iCub and Shadow hand kinematics ((C) Shadow Robot Company 2014). The kinematics of the iCub hand thumb has four DoFs and is under-actuated, while the kinematics of the Shadow hand thumb has five DoFs and is fully actuated.

using the conventional equations [158] and the Denavit-Hartenberg (DH) parameters of the two hands. The origin frame of the kinematic chains of the thumbs was set to the centre of the dorsum of each respective robotic hand, so that the calculated

thumb positions for both hands are expressed in the dorsum reference frame. The DH parameters of the Shadow Hand were obtained from the work of Cui et al. [159], and the joint ranges were taken from the Shadow Hand Technical Specifications [70]. The origin frame of the Shadow Hand was translated from the origin reference point of 280 mm on the Z axis (depth). The DH parameters and the joint values of the iCub hand were obtained from the iCub online manual [160]. The origin frame was translated from the reference point of -4.3 mm, -3.3 mm and -19.1 mm on the X (length), Y (width) and Z axes respectively. The reachability space of the iCub hand was calculated using all available joint values using one degree step. The reachability space of the Shadow hand was evaluated using intervals of 5 degrees due to the high computational complexity of the calculation. The total number of data points generated from the calculations are of comparable size: 670,761 and 504,735 for the iCub and the Shadow hand respectively. The size of the iCub hand is slightly smaller than an average human hand. Therefore, it was required to isotropically scale up the data calculated from the iCub forward kinematics by a 1.23 factor. This coefficient was calculated by dividing the length of the iCub hand (165 mm) by the mean length of male hand (197.1 mm according to [27]) and inverting the resulting scaling matrix.

The human hand is a *biological gripper* and its reachability space can be described as a function of many subject dependent variables, such as elasticity of tendons, cartilaginous flexibility of the articulations, length of the bones, tendons and muscles and more. Therefore, the reachability space of the human thumb cannot be formulated as an equation (closed form) in general. It is possible to calculate the reachability space of the human thumb using a specific kinematic model from literature. However, in the analysis of this chapter, the reachability space was derived from the thumb motion data of the GRASP database to avoid any bias and inaccuracies imposed by any kinematic model. Human data from the Primate Grasp posture class was excluded from this analysis, as the thumb is not used in prehensile movements by definition of this posture.

The human and robot data were divided in three groups, one for each plane of the 3D space, in order to simplify the comparison of the reachability spaces. The human and robot reachability spaces were expressed as point clouds. For each plane in 3D space, the point clouds were fitted with the minimum convex hull, which represents the minimum cluster of the point cloud. On comparison, a regular convex hull can be of any dimension. The approximation to convex hulls is used, as the point cloud for human and robot reachability spaces is nearly convex. As a result, there were obtained three convex hulls for each human posture class (Power, Precision and Key Grasp), and two convex hulls for each robot hand. The two robot hands were compared with the human data separately, as can be seen in Figure 4.7. The comparison was based on their geometrical overlap with the reachability spaces of the human thumb and it was calculated numerically. Additionally, the volumes of 3D reachability spaces of the human thumb for the Power and Precision Grasp postures and the volumes of the robotic thumb reachability spaces were compared to the mean palm area. The mean palm area was calculated as *palm length times palm width*, using data collected in [27] from more than thousand male subjects. The mean palm length is $110.5mm$ and the mean palm width is $95.3mm$. The volume for each 3D reachability space was calculated on the minimum convex hull enveloping the point clouds in 3D space. The Key Grasp was excluded from the analysis, as the thumb is not opposing the other fingers by definition and it would be inappropriate to compare it to the area of the palm. This comparison is defined as postural spanning ratio and was calculated as follows:

$$\frac{V_h}{A_p} \tag{4.5}$$

Where A_p is the mean palm area and V_h is the volume of the convex hull of any reachability space used in the comparison (Table 4.3).

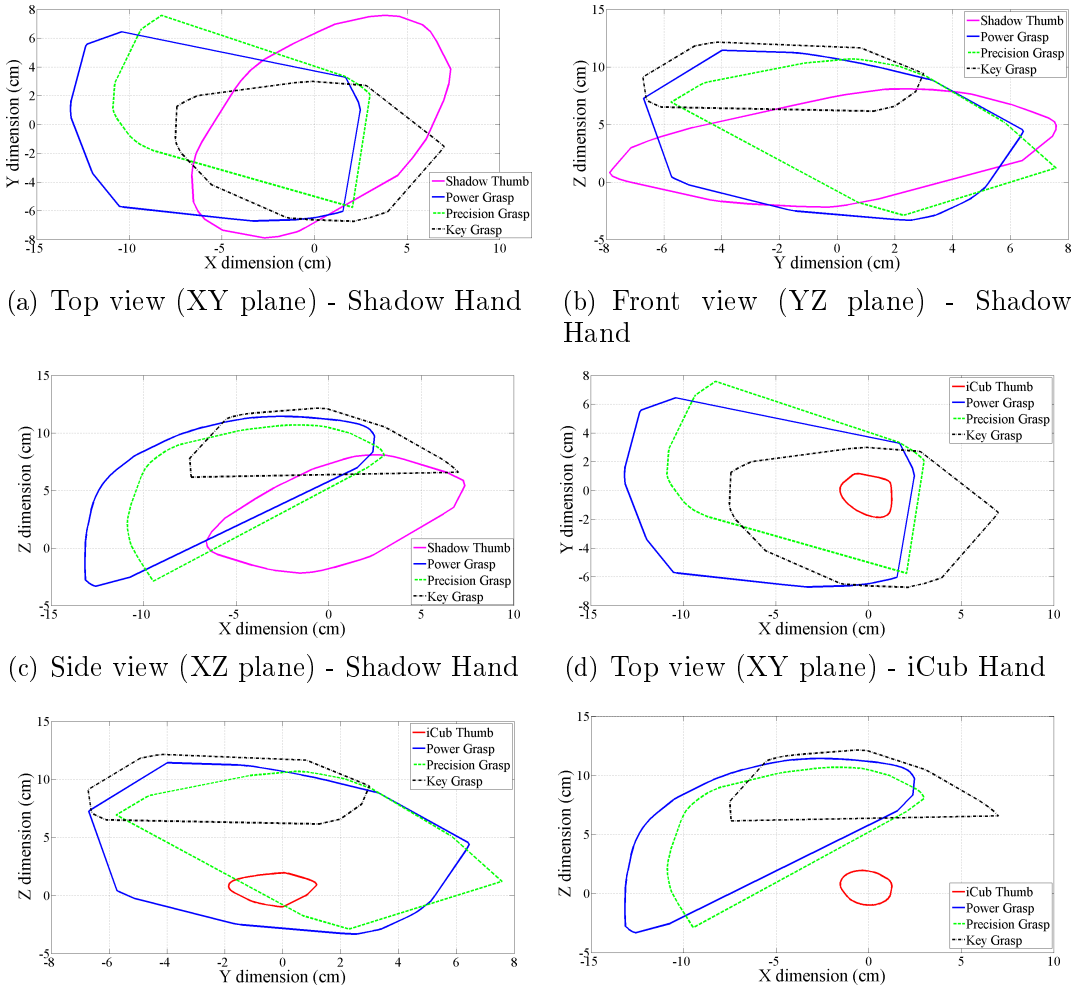
If the convex hull of a robot hand overlapped the convex hull of human posture class, the

area of the overlap was calculated using the Sutherland-Hodgman algorithm [161]. This algorithm, widely used in computer graphics, calculates the intersection between two polygons - a clipping polygon and a subject polygon. In this case, the clipping polygon was the robot hand convex hull and the subject polygon was a human posture class convex hull. The intersection area was calculated for each posture class and for each plane across the two robot hands. An intersection area was expressed as a percentage of coverage, by dividing it by the corresponding class area of a human posture. Finally, the total percentage of coverage for each posture class was calculated as the average across the three planes.

4.4.3 Comparison of Human-Robot Reachability Spaces

This part shows the results of the numerical comparison to evaluate whether the reachability space of selected robotic thumbs is evenly spanning across the reachability space of the human thumb for each posture class. Figure 4.7 shows the reachability spaces expressed as convex hulls for each human posture class and robot thumb across the three planes. From this figure, it can be observed that the reachability spaces of the robot thumbs are not evenly covering the reachability spaces of the human thumb. As the Shadow Hand is more articulated than the iCub hand, a larger space of the human data is covered, but still the space does not evenly correspond to the human reachability space for all posture classes. Both hands do not perform well when covering the motion in the side view (XZ plane), as shown in Figure 4.7(c) and Figure 4.7(f).

In addition, it can be observed that the Key Grasp is the most under-represented category of grasps in robotic thumbs. It is covered only by 35.7% and 2% of the Shadow and iCub hand reachability space respectively. The best represented class is the Precision Grasp, covered by 46.65% and 3.15% of the reachability spaces of the Shadows and iCub hand respectively. The results of the numerical comparisons for each grasping



(a) Top view (XY plane) - Shadow Hand

(b) Front view (YZ plane) - Shadow Hand

(c) Side view (XZ plane) - Shadow Hand

(d) Top view (XY plane) - iCub Hand

(e) Front view (YZ plane) - iCub Hand

(f) Side view (XZ plane) - iCub Hand

Figure 4.7: The figures compare the human thumb reachability space, derived from the GRASP database captures and organised by posture class, with the Shadow thumb reachability space (continuous magenta line) and iCub thumb reachability space (continuous red line) calculated from their respective forward kinematics. The polygons represent the minimum convex hull enveloping the data samples. The Primate Grasp posture class has been excluded from the calculation as the thumb is not relevant for grasping by definition.

class are reported in Table 4.2. These findings can be explained by the fact that robotic grasping is traditionally focussed on contact-point based precision grasping. These findings can be explained by the fact that robotic grasping is traditionally focussed on contact-point based precision grasping. Hence, the design principles of robotic hands

are more adequately representing this category of grasps.

Table 4.2: Summary of the percentage of overlaps between robot thumb reachability space and human thumb reachability space divided by posture class.

Grasping Class	Shadow Hand	iCub Hand
Power Grasp	41.02%	2.63%
Precision Grasp	46.65%	3.15%
Key Grasp	35.73%	2%

4.5 Comparison of Human-Robot Postural Spanning Ratios

The Shadow hand is considered to be one of the most dexterous robotic hands available, but the analysis shows that the representation of the reachability space of the human thumb is below 50% for all posture classes. Additionally, the volumes of the reachability spaces of the robotic thumbs and of the human thumb were compared alongside with the spanning ratio between the volumes and mean area of the male palm. The results in Table 4.3 show that the surface of the palm, covered by human Power Grasp, is larger than the surface covered by the Shadow Hand thumb. The spanning ratio with the mean palm area was larger for the human Power Grasp, than for the Shadow Hand's thumb. However, the reachability space for the Precision Grasp follows an opposite pattern. In this case the Shadow Hand thumb spans a larger area than the human reachability space. Overall, the spanning ratio of the Shadow Hand is much inferior to the human thumb.

In general, the Precision Grasp is the class of posture best rendered by present state-of-the-art robotic thumb kinematics, while the Key Grasp is the poorest represented class. This suggests that some grasp affordances can be easier implemented in robotics since they leverage on a hand posture that is easier to replicate. In these cases, the thumb kinematics can replicate a very similar posture to the human thumb. Such

Table 4.3: Summary of the volumes of the 3D convex hull enveloping the reachability spaces and the spanning ratio between area of the palm and volumes of the hulls.

3D Convex Hull	Volume (mm^3)	Spanning Ratio
Power Grasp	956.1	0.0916
Precision Grasp	627.4	0.0596
Human Overall	1157.1	0.1099
Shadow Hand	757.6	0.0719
iCub Hand	11.3	0.0011

feature simplifies the approach strategy when aligning the palm to the object. Instead other grasp affordances might need an ad-hoc approaching strategy to obtain results comparable to human grasping, since the affordance requires a hand posture poorly represented by state-of-the-art robotic thumbs and a more careful alignment of the palm to the object.

Humans are the only primates that are able to manipulate objects in a very complex way, and most of everyday objects are tailored to our needs. Therefore, it is likely that robotic thumbs will be better poised to manipulate objects that are designed for human hands if they would capture the essential morphological features of human thumb. In this case, it is important to highlight that the kinematics of future robotic grippers should consider extending the reachability space of the thumb by opposing the whole trapezium-trapezoid-scaphoid complex (foldable part of the palm that belongs to the thumb), as it is in humans, rather than just opposing the base of the thumb metacarpal.

However, it is worth to note that it is possible to perform suitable grasps without using the thumb, as is the case for many primates. Such approach is limited to the use of one posture class only (Primate Grasp), hence reducing the dexterity of manipulation.

4.6 Conclusions

The prominent role of the thumb in prehensile grasping is detailed in this chapter. Statistical analysis and quantification of the motion variability of human fingers is performed on a dataset of 31 different postures used on various objects. Additionally, a classification of the hand posture that is based on the thumb is defined. The prehensile reachability space of the human thumb is then compared with the reachability space of two robotic hands, the iCub hand and the Shadow hand, to understand how well state-of-the-art robotic thumbs render different posture classes. The reachability space of robotic thumbs unevenly spans the human thumb reachability space, as some posture classes, such as the Precision Grasp class, are better rendered by others. It can be concluded that state-of-the-art robotic thumbs can limit the selection of implementable grasp affordances, as some finger postures might be difficult to replicate on a robotic hand. The approach strategy must take into account such limitations when positioning the palm near the object by, for example, adjusting the hand placement to use an alternative grasp posture.

The main findings of the chapter are listed below:

1. The analysis of motion variability demonstrated that the motion of the thumb is leading the motion of other fingers in grasping; and that the thumb position depends only on the specific grasp posture configuration, independently from the manipulated object.

2. A human grasp posture classification based on the position and use of the thumb is outlined.

3. The comparison between human and robotic thumb reachability spaces show that Precision Grasp is the best rendered posture by robotic thumbs, as the overlaps

between the two reachability spaces is maximum. For the same reason, Key Grasp is the worst represented posture.

4. The comparison of the spanning ratios between the robotic and human reachability spaces and the mean area of the human hand demonstrated that the reachability space of state-of-the-art robotic thumbs covers a larger area than human Precision Grasp, but a smaller area than human Power Grasp. This confirms that human Precision Grasp is better rendered by the state-of-the-art robotic thumbs, and such bias can influence and limit the selection of implementable grasp affordances.

5. Suggestions on an improved kinematic designed of robotic thumbs are provided. It is advised to actuate the whole trapezium-trapezoid-scaphoid complex rather than just the base of the thumb metacarpal, to improve the spanning ratio.

Chapter 5

An Architecture for Action-Focussed Grasp Affordances

5.1 Introduction

This chapter proposes a modular framework which models the main aspects of a grasp affordance as individual separable components. It also proposes a trajectory generator, controlling the motion of a robotic hand, that is based on the results of Chapter 3. The main target of the approach proposed in this chapter is to reduce the dependency, of robotic grasping, from modelling the properties of an object. As alternative, more significance is given to fulfil an assessment criterion which evaluates the effectiveness of a grasp control strategy. This approach can be applied to an arbitrary action performed with any graspable object. In this chapter, the validity of this approach is assessed through a feasibility study in a simulated environment. The position of the end effector is learned from the algorithm; therefore, it is not strictly required to execute the algorithm on a robot with human-like arm kinematics. However, the reachability space of

the robotic arm has to be comparable with the one of the human arm. As the proposed Grasp Affordance Framework enables a robot to perform an action using a graspable object, it is possible to formally assess the quality of execution of the desired task with a performance metric. A complex task, such as the assembly of a Lego-style object, does not fit within this framework, as the sequence of a task cannot be unequivocally assessed using a single performance metric. Complex tasks should to be divided in smaller and easier to assess actions. For simplicity of explanation, it is assumed that an object is already recognised and segmented in individual primitive graspable shapes.

This chapter is a feasibility study of the proposed architecture. It is explained how the Grasp Affordance Framework can be structured and how it can be implemented. In addition, the limitations of this approach are identified. The study focus only on the reaching part of the framework, a suitable grasping controller is described in Chapter 6.

5.1.1 Motivations

Traditional approaches to manipulation control require a hand/arm motion controller and a learning algorithm which tunes it. The learning algorithms used can be Gaussian mixture models [111], Markov decision process [162], or many others [112]. The common problem of this methodology is that the resulting system is tailored on a narrow and specific application [116, 126]. It is difficult to scale over a wider range of objects, because the grasping action is mostly the target of learning. The approach proposed in this chapter changes the problem formulation. Performing an action to fulfil a minimum level of quality is the target of learning, and the precise details of grasping are overlooked. As long as a grasp action meets the set quality criterion, it is unimportant which specific object is used or how it is grasped. To obtain such flexibility, the system should be modular, so that different learning algorithms or controllers can be used to better fit

different actions.

The proposed approach uses human demonstrations as a foundation to derive control models to generate a trajectory for a grasp action execution. The models derived in Chapter 3 are used for robotic implementation in this chapter. Demonstrations performed by observing a human teacher require the resolution of the correspondence problem [113]. The problem was addressed by learning the evolution of a representation of the end effector over time. Such representation is the centroid of the end effector, which is independent from specific hand or arm morphologies. The pattern of evolution of the centroid corresponds to the action executed, and transcends the specific kinematic details of the teacher or the learner. An alternative approach is to use Dynamic Motion Primitives (DMPs) [163] that mathematically model the actions of a human teacher and map them to the robot intrinsic space - the robot's own embodiment. DMPs are designed to learn skills, such as swinging a tennis racket. This approach can be an alternative solution to produce end effector trajectories from demonstrations. However DMPs do not take into account the complications of grasping due to the control of the fingers; those are addressed in this approach.

5.1.2 Summary of Contributions

The contributions of this chapter are the following:

1. The general architecture of a grasp affordance framework is presented and grasp affordances are re-defined in terms of object-hand relationships for the purpose of robotic manipulation.

2. A human-inspired trajectory generation algorithm of approaching to grasp is shown, and two different control hypothesis are formulated.

3. The feasibility of the proposed control algorithm and models are evaluated on a human-like robot arm, and their performances are discussed and compared with a state-of-the-art planner in a simulated environment.

The rest of the chapter is organised as follows. Section 5.2.1 presents the general architecture of the system and the main concepts. Section 5.2.2 introduces the algorithm of the trajectory generator of approaching that translates hand-to-object distance profiles in Cartesian space of the end effector. Sections 5.3 and 5.4 describe the experimental environment and results. Section 5.5 discusses the results, proposes future extensions of the algorithm and draws the conclusions.

5.2 Description of the System

5.2.1 Overview of the Architecture

Prior to describing the proposed grasp affordance trajectory generator, it is necessary to provide a general overview of the grasp affordance system. The proposed system is summarised in Figure 5.1. Its purpose is to define a suitable hand trajectory and grasp strategy for gripping an object and use it to perform a simple action. The outcome of a desired action must be measurable so that the performance of different trials can be evaluated. It is assumed that the same area of the object affords different successful approaching and grasping strategies. Each successful grasp can achieve a different performance when executing the desired action. For example, hammering can exert higher loads if a mallet is held by a handle with the whole hand, rather than just with the fingertips.

Further on, notations used in this chapter are defined. A *robot state* can be encoded as a set of variables representing the internal state of the robot relevant to grasping,

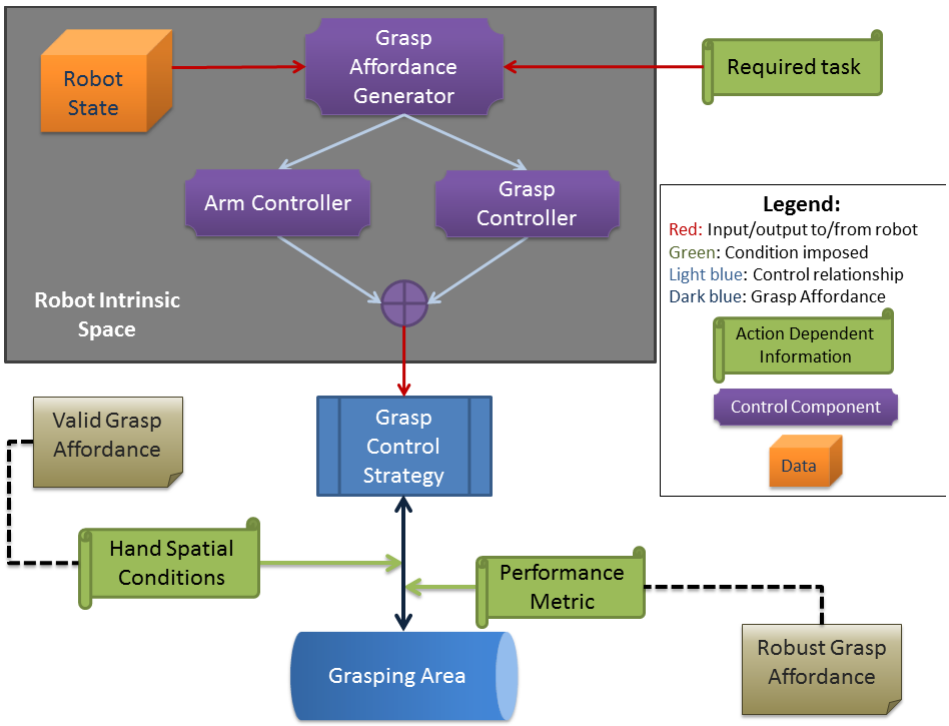


Figure 5.1: Architecture of the proposed Grasp Affordance System. Purple elements are components of the system, here implemented in software. The grey area identifies the elements that directly interact with the robot. The arm and grasp controller can be as simple as the inverse kinematics and as complex as an attractor-based controller, or expressed in closed form, such as a Grasp Jacobian, or be free form as a mechanical-based synergistic hand.

such as joint positions, and it is specific to the embodiment of the robot. A *grasp control strategy (or grasp strategy)* is the evolution of a robot state over time and it aims at gripping a part of an object. A *grasping area* is a section of the object that is considered graspable with the available end effector. In addition, a grasping area can be modelled as a 3D geometric primitive, such as a solid of revolution. A single object can be segmented in multiple grasping areas; an example is given in Figure 5.2

It is computationally easier to produce a grasping control strategy for multiple simple geometrical shapes compared to a complex geometrical structure of a complete object, due to well-known properties of basic solids. Another advantage is the possibility to

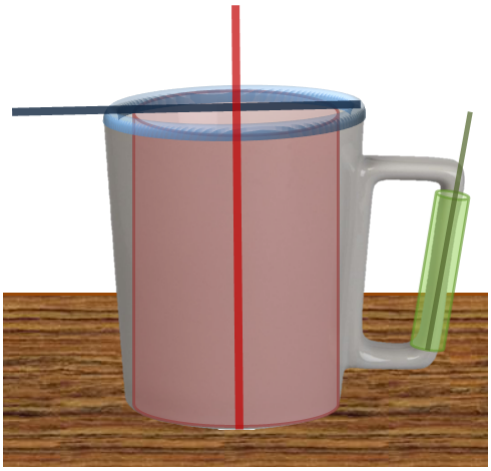


Figure 5.2: Possible segmentation of a mug in three primitive shapes representing *Grasping Areas* suitable for grasping. Each area is a simple rotational shape and it has a principal axis along its longest dimension. In the Figure, a principal axis is a line of the same colour of its grasping area. The red and green cylinders are grasping areas segmented from the body and handle of the mug. The blue toroid is the grasping area of the mug's border.

re-use successful grasping strategies for wide range of objects with grasping areas of the same shape, in a similar way the humans already do [48]. For instance, a grasping strategy that can successfully grasp a handle, rendered as thin cylinder, could also grasp a segment of any other object approximated with same shape.

In this chapter, for the purpose of robotic manipulation, a grasp affordance is modelled as *the association between a grasp strategy and a grasping area.* Therefore, a grasp affordance is characterised in terms of associations between the properties of the area and those of the hand. This concept is illustrated in Figure 5.3. By imposing conditions on such relations, it is possible to assess whether a grasp strategy grips on the given area satisfactorily. Such grasp strategy, and its specific combination and evolution of variables that generated it, is defined as a *valid grasp affordance.* For example, the centre of the hand and the centroid of the grasping area can be associated by their relative distance and orientation. It can be imposed that a valid hand trajectory must

position the hand at a fixed distance and orientation with respect to the grasping area's principal axis to allow the fingers to envelop the object correctly and guarantee a grasp. If there is a grasp strategy able to fulfil all the conditions, then that is considered a valid grasp affordance for the area of interest.

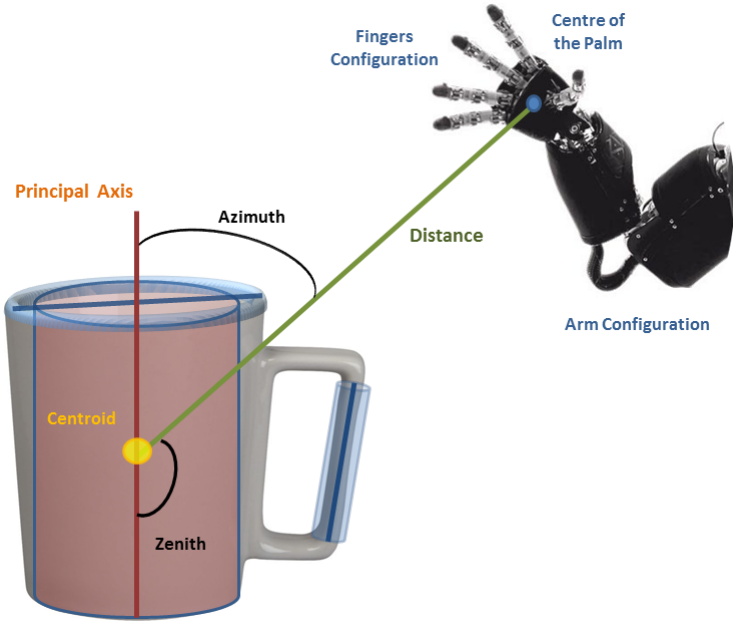


Figure 5.3: A Grasp affordance is a set of numerical associations between properties of the robotic hand/arm (dark blue) and properties of a grasping area (orange). A Grasp Affordance is relative to a single grasping area (red), other grasping areas (blue) can be linked to the robot using similar associations.

The framework allows the robot hand to initiate a grasp action in parallel with the approaching motion and before it enters in contact with the object, as observed in humans [151]. A grasp controller suitable for a parallel execution is described in detail in Chapter 6, but it is not presented in this study since this chapter only assesses the feasibility of the approaching part.

It is important to underline that the same grasping area can be gripped in many different ways which satisfy the definition of valid grasp affordance. For instance, the handle of a mug can be grasped with a pinch grip or a full power grasp and both postures are valid

grasp affordances. In some cases, a grasp strategy can be successful, but the generated grip might be too weak to execute any action than pick and place. The subset of successful, but ineffective, grasps is defined as *unfit grasps*. Conversely, *solid grasps* are defined as appropriate for the desired purpose. To distinguish a valid grasp affordance that produced unfit or solid grasps, the concept of *robust grasp affordance* is defined. A valid grasp affordance is defined as robust if a predetermined action is successfully executed satisfying a minimum performance threshold. In other terms, a valid grasp affordance with a solid grasp is defined as a robust grasp affordance. The performance of an execution has to be is evaluated in closed form, using an equation, and is action specific. For example, consider the hammering task implemented in Chapter 3. A valid grasp affordance is considered robust if it is able to produce a minimum amount of force when hammering on a point. The force threshold has to be consistently exceeded for a set number of trials to qualify a valid grasp affordance as robust.

It is fundamental to specify in advance, which task and performance metric should be used. If the above-mentioned conditions are not defined, it cannot be established what is a robust grasp affordance, since any valid grasp affordance could be suitable.

5.2.2 A Grasp Trajectory Generator for Approaching to Hammer

An approaching controller, to guide a robotic hand to a valid grasp position, is described in this section. The proposed algorithm is a trajectory generator based on the identified approaching patterns of humans studied in Chapter 3. It was found that there is a common structure in approaching an object to hammer. Two hypotheses defining the approaching pattern in closed form were provided: an approximate model and a maximally reliable one. Those hypotheses model the evolution over time of the approach speed of the hand to the object. The speed describes a hand-to-object relationship, but a robot hand requires a trajectory of 3D points to be controlled. Hence it is required

to convert the evolution of the speed in a sequence of 3D points. In this section an algorithm that converts the profile of the speed of the distance to 3D points used in Cartesian control is presented. Each point of the profile is related to the current hand position and converted to a new 3D point for the next position of the end effector. Further on, the arm controller can reach for the desired point using any form of Cartesian controller. The algorithm can be used with a different model of approaching which renders differently the same hand-object relationship or a completely different property. Since this is a feasibility study, the inverse kinematics of the selected controller is used to calculate the joint positions from a Cartesian point. Alternatively, an attractor-based control system [164] is an equally suitable and computationally inexpensive option.

In the proposed implementation of the algorithm, the robotic hand is approaching the object until the controller reaches a given point within a set distance. If needed, a different termination criterion can be chosen, such as requesting the normal of the palm to be orthogonal with the principal axis of the object. Upon successful termination of the trajectory generator algorithm, the robot hand is expected to be positioned appropriately for finalising the grasp. This corresponds to the beginning of the Stage 4 observed in human studies of approaching. The controller relies on one of the two hypotheses to compute the approaching trajectory of the hand. The first hypothesis - maximally reliable hypothesis - is a model that is based mostly on non-polynomial equations which closely render the original human models. The second hypothesis, the approximated hypothesis, uses only 2nd order polynomials to guide the hand, and is the best approximation of the original human motion. The second hypothesis aims to maximise the simplicity of the model of approaching at the cost of reduced accuracy. The first hypothesis privileges precision over simplicity. Each hypothesis is the combination of three models defined for each approaching stage in Chapter 3. The second model is common for the two hypotheses. To switch between a current model and the next one, the position of the end effector at the last valid point of the model is linearly

interpolated to the first valid point of the subsequent model.

Algorithm 1 describes the conversion of the speed of the distance to a trajectory in Cartesian coordinates in detail. Its main parts are now explained step by step. Initially, the algorithm derives the direction cosines manipulating the formula of the conversion between Euler angles and quaternions:

$$\frac{\alpha}{2} = \texttt{arcos}(q_w) \tag{5.1}$$

$$\cos(\beta_x) = \frac{q_x}{\texttt{sin}(\frac{\alpha}{2})} \tag{5.2}$$

$$\cos(\beta_y) = \frac{q_y}{\texttt{sin}(\frac{\alpha}{2})} \tag{5.3}$$

$$\cos(\beta_z) = \frac{q_z}{\texttt{sin}(\frac{\alpha}{2})} \tag{5.4}$$

Where $[q_x\ q_y\ q_z\ q_w]$ is the quaternion representing the current robot orientation and $\beta_x,\ \beta_y,\ \beta_z$ are the Euler angles across the X, Y, Z axis of rotation. Since the human derived models operate on non-dimensional quantities, the norms of the Cartesian coordinates of the end effector of the robot and of the position of the object are normalised to the unity. At this point, it is possible to calculate the hand-object distance and to derive the speed of approach as follows:

$$d_t = \|\overline{h}\| - \|\overline{o}\| \tag{5.5}$$

Where $\|\overline{h}\|$ and $\|\overline{o}\|$ are positions of the end effector and the object normalised to the unity and expressed in the robot base reference frame. The above calculation corresponds to Equation 3.2.2 used to derive the hand-object distance from human data in Chapter 3. Further on, the speed of approach can be obtained from the specific

Algorithm 1: Control algorithm of the grasp affordance trajectory generator. This algorithm converts a hand-object distance relationship, encoded in a model of approaching learned by human demonstrations, in a 3D end effector position.

Data:
eePos, eeQuaternion: position and orientation in quaternions or the robot's centre of the gripper, in robot reference frame coordinates
obPos: the position of the centroid of the reached object, in robot reference frame coordinates

```
begin
    while !stopConditionMet(eePos) do
        eePos ← robot.forwardKinematics(position);
        eeQuaternion ← robot.forwardKinematics(quaternion);
        /* Obtain direction cosines                          */
        alphaOver2 ← arcos(eeQuaternion[w]);
        dirCos[X] ← eeQuaternion[X]/ sin(alphaOver2);
        dirCos[Y] ← eeQuaternion[Y]/ sin(alphaOver2);
        dirCos[Z] ← eeQuaternion[Z]/ sin(alphaOver2);
        obPos ← env.getObject(position);
        /* Normalise positions to unity                      */
        eePos ← normalise(eePos);
        obPos ← normalise(obPos);
        /* Convert positions to norms and distance           */
        eeNorm ← norm(eePos);
        obNorm ← norm(obPos);
        distance ← eeNorm-obNorm;
        /* Calculate next speed from specific approaching
           model                                             */
        newSpeed ← gaffordance.calculatePolicy();
        newDistance ← newSpeed +distance;
        /* Split distance into XYZ components and add it to
           current position                                  */
        eePos[X]-=newDistance * dirCos[X];
        eePos[Y]-=newDistance * dirCos[Y];
        eePos[Z]+=newDistance * dirCos[Z];
        /* Convert back to metric distance to allow robot
           control                                           */
        eePos ← convertToMetricDomain(eePos);
        reacher.moveEEto(eePos);
    end
end
```

model of approaching embedded in the trajectory planner. From the speed and the distance it is possible to obtain the expected next hand-to-object distance:

$$d_{t+1} = d_t + s_t \tag{5.6}$$

Where d_t is the current hand-object distance, calculated from the hand and object positions, s_t is the speed derived from the models and d_{t+1} is the new distance which it is expected to be reached at the next iteration. Since the object is static, it is possible to derive the next hand position from the distance. The already calculated direction cosines can be used to convert the norm back to Cartesian coordinates used for the arm controller, as follows:

$$h_x = d_{t+1}\, cos(\beta_x) \tag{5.7}$$

$$h_y = d_{t+1}\, cos(\beta_y) \tag{5.8}$$

$$h_z = d_{t+1}\, cos(\beta_z) \tag{5.9}$$

Where h_x, h_y and h_z are the three dimensional components of the hand position expressed in robot base reference frame. Finally, functions of Algorithm 1 (Trajectory Generator) are now explained. The *gaffordance.calculatePolicy()* is the function that calculates a speed of approach from a human derived model. Function *stopConditionMet()* determines a suitable hand spatial conditions to terminate the algorithm. In this case it is set as a minimum distance between the hand and the object centroids. The *convertToMetricDomain()* function brings the end effector coordinates back into a dimensional domain (cm), while *reacher.moveEEto(eePos)* asks the arm controller to move the end effector to the next desired location *eePos*.

5.3 Experimental Setup

The robotic platform selected for this feasibility study is the Baxter robot, commercially available from Rethink Robotics. Baxter is a fully compliant robot, designed for interaction and close cooperation with humans. Baxter has a human-like arm kinematics, composed of 9 rotational joints [165]. The maximum length reachable by its arm is about 1.3 meters and it is equipped with a parallel gripper which opens and closes [165]. The kinematic structure of the robot slightly differs from the human one, since the wrist and shoulder joints have Cartesian offsets, and the range of some joints does not reflect the original human anatomy [165]. Despite this, the range of motions and the workspace covered by the robot's arm is similar to the one covered by human arms.

The robot is controlled with Linux-based Robot Operating System (ROS) [166], which is the state of the art in robotic software control. The models of Chapter 3, derived from human data, and the algorithm described in this chapter, were implemented in C++ taking full advantage of existing ROS tools and the GNU Debugger. This specific implementation heavily relies on floating point operations, in IEEE-754 format, to perform the calculations. To reduce noise and inaccuracies normally introduced by floating point arithmetic, all the sums and subtractions are performed with the Kahan Summation Algorithm [167]. The results of the simulation were assessed in simulation using Gazebo - the standard ROS 3D simulator based on Ogre physics engine. As a baseline for comparison, an optimised version of the original Rapidly-Exploring Random Trees planning algorithm [168], the RRT-connect [169], was used to generate motion plans for the end effector. This planner was selected, as it is one of the most popular algorithms used for robotic motion planning. The algorithm is provided by the Open Motion Planning Library (OMPL) [170] which is the core component of the MoveIt! [171] robotic planning framework. The framework was used to generate the motion plans via its graphical user interface, a plug-in for the graphical robot monitor tool Rviz.

As mentioned before, any algorithm that is able to convert end effector position in joint angles can be used as arm controller. In this work, Baxter's built in Inverse Kinematics was used. However, an alternative algorithm based on field attractors, such as the Passive Motion Paradigms [121], can be employed too to save on computational costs. In this chapter, the performance of the affordance generator in driving the arm is assessed. Hence, the desired task is a simple reaching task that does not involve any manipulation. For this reason the grasping area was simplified to a cube of 4 cm^2 from the centroid of the object. Since no manipulation is required at this stage, a performance metric cannot be chosen. Instead, the following spatial conditions were imposed:

$$5.5 \geq \|\overline{h}\| - \|\overline{o}\| \geq 4 \tag{5.10}$$

Where 4 is the distance between the object centroid and its edge expressed in centimetres. $\|\overline{h}\|$ and $\|\overline{o}\|$ are the norm values of the end effector and object positions in the robot base reference frame. As it can be observed, this condition is imposed on the approach distance of a grasp affordance relationship, as this is the most important part to be assessed. The approach distance is fundamental to be rendered correctly, as a hand placed too far away from the object cannot grasp it. Also certain conditions, like a spherical geometry of an object or a target point placed on the surface instead of the centroid, might relax the requirements on the end effector orientation. Following the methodology of human experiments, the robot hand position is modelled as the centroid between the two fingers of the parallel gripper.

The experimental set-up is designed to match human demonstrations described in Chapter 3. This is required to ensure consistency between the human demonstrations and robotic implementation. The initial position of a robot arm is comparable to the starting position of the human experiments. The robot arm is straight and parallel to the

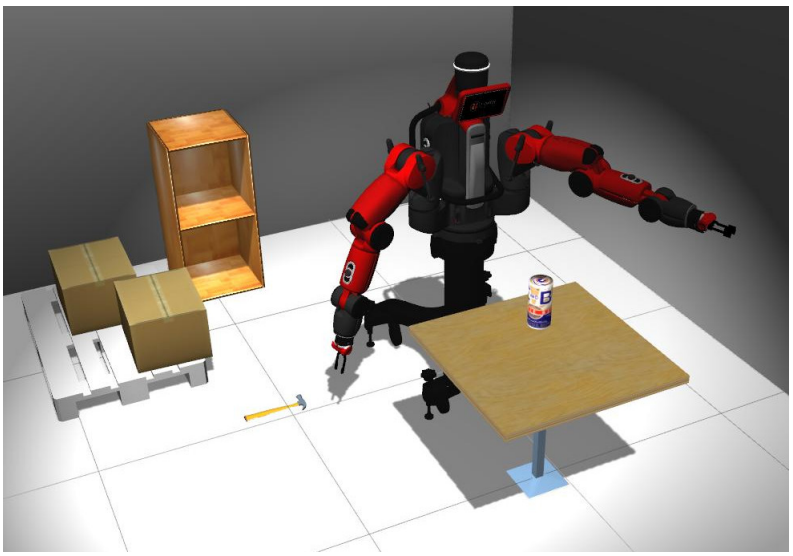


Figure 5.4: A simulated Baxter robot is employed to assess the feasibility of the proposed grasp affordance system. The set-up replicates the main assumptions imposed when recording human demonstrations for the underlying control models. Specifically, the arm posture and the object position are similar to the human counterpart.

ground and the parallel gripper open. The object was placed on the table within an area of the robot's workspace reachable by both arms, as it was performed for human experimentations. To assess the feasibility of the system, the grasping area of the object was approximated to its centroid. The complete set-up is shown in Figure 5.4.

5.4 Results and Comparison

To assess the feasibility of the implementation and to compare the performance of the proposed method with the state of the art, the object was placed in four different locations that cover different areas of the workspace. The first three locations are within the reachability space of both arms of the robot, and the last is at reachable by the left hand only. The coordinates of the four positions are listed in Table 5.1. The core element of the grasp affordance trajectory generator is the normalised profile of

the hand-to-object approach distance speed, evaluated from human studies. As such, the experimental data should be converted to the same format to conduct a consistent analysis. For this reason, the trajectory of the end effector, produced by RRT-connect, was related to the position of an object. Then, the normalised speed of this relationship was calculated and the length of each trail was matched with the length of the human demonstrations. As the speed data calculated from RRT plan is noisy, a median filter of size seven was applied to the trials. As mentioned in Section 5.2.2, the output of the trajectory generator was created using two different models of approaching: a simple approximated one and a complex maximally reliable one. The output of both models is the speed of approaching matching the format of human demonstrations. Figure 5.5 shows the outputs of the three examined techniques for the four positions. It can be observed that RRT-connect consistently led to the collision with the table when executing a motion plan. The RRT algorithm allows to include constraints to avoid collisions but it has been preferred to disable the collisions completely as the aim of this study is to compare the two algorithms in their simplest form. A similar behaviour was also observed with the other planners offered by the MoveIt! framework.

Table 5.1: The same object was positioned in four locations expressed in the robot reference frame. *Low* indicates the wrist, *high* indicates the head, *outer* indicates the side, *in the middle* indicates the middle of the trunk. *Far* and *close* are measured in terms of arm extension.

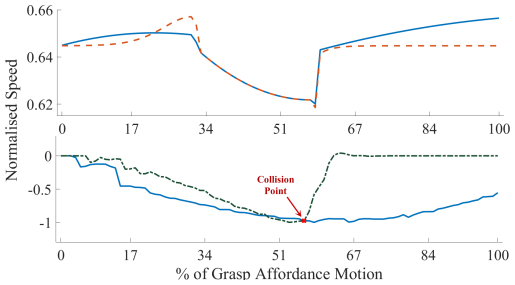
Object Position	**Coordinates in Robot Frame**		
	X (cm)	**Y (cm)**	**Z (cm)**
Low and close	50.11	1.57	-5.89
High and far	90.81	4.28	57.86
In the middle	58.57	-0.49	35.96
Low and outer	66.5	64.84	2.64

RTT-connect and the maximally reliable model generated a trajectory able to successfully reach the desired point in all cases. However, the output of the approximated model was not stable. The evaluation has to be therefore based on different factors.

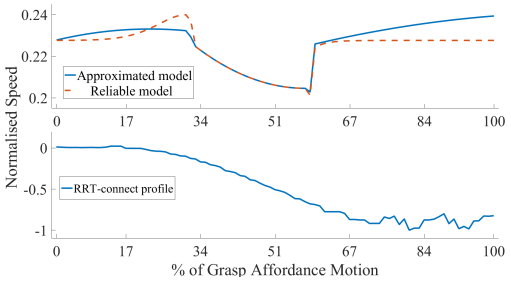
The collision avoidance is intrinsically guaranteed by the grasp affordance trajectory generator. No further constraints are required, as the action execution was derived from human demonstrations, while the RRT planner requires additional conditions and complexity in this case. However, RRT can handle any reaching situation, if the collision avoidance conditions are specified correctly. The trajectory generator, instead, is tailored to one action.

The trajectory generator is composed of different models, in closed form, combined in order. Each model can be changed as long as its purpose and functionality in the sequence is preserved. RRT, in contrast, provides a monolithic solution based on a single algorithm. Hence, it cannot be separated in stages, and it is less flexible than the approach suggested here. Since the trajectory generator is a juxtaposition of models in closed form, every equation can be easily and efficiently calculated within milliseconds. RRT, instead, relies on a complex tree exploration based algorithm, which offers less optimisation capability and higher computational complexity of the order of the seconds. The motion produced by the trajectory generator is less noisy, while the RRT planner requires filtering. However, a different implementation of RRT might mitigate this problem. It can be concluded that the RRT planer is a complex and computationally-expensive approach to generate a trajectory in any situation, but it is unable to provide any further customisation for realising grasp affordances. The proposed trajectory generator, instead, is a lightweight algorithm that offers more freedom of configuration and addresses affordance-specific problems. However, it requires an extra complexity to be able to generalise to any action in any condition.

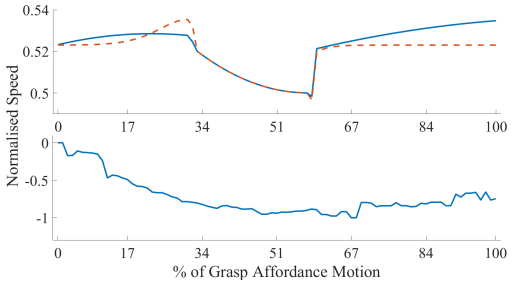
Figures 5.5 demonstrates that the best approach of control is a mixture of the two different control hypotheses used to generate trajectories. The approximated hypothesis has a smoother and simpler profile in the beginning, but the speed profile diverges at the end. The maximally reliable approach, instead, converges in the last stage. However,



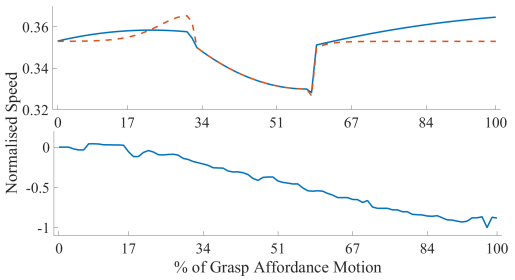
(a) Speed profiles for object in position 1



(b) Speed profiles for object in position 2



(c) Speed profiles for object in position 3



(d) Speed profiles for object in position 4

Figure 5.5: Speed profiles of two trajectory generator hypotheses (upper graph) and the RRT-connect planner (lower graph). Figure 5.5(a) shows an execution of RRT-connect colliding to an obstacle (dash-dotted line). The solid line in the upper graph represents the approximated speed profile, while the dashed line represents the maximally reliable profile. The trajectory generator relies on three models, the second is common between the two hypotheses.

this hypothesis exhibits a peak before a sharp drop in speed in the first part of the control policy. Such behaviour is not desirable, as it could generate unwanted vibrations. Therefore, the best approach employs an approximated model for the initial stage, and a maximally reliable for the last (the second stage is common for both hypothesis). The modular structure of the control strategy allows combining different models together.

5.5 Discussion and Conclusions

In this chapter, a modular architecture for a grasp affordance controller was presented and grasp affordances in robotics were modelled in terms of robot-object associations. An algorithm for generating hand trajectories from two control hypotheses learned from human demonstration was proposed. A feasibility study of this system was carried out in a simulated environment, and the system was assessed against a state of the art planner, the Rapidly Exploring Random Tree connect (RRT) algorithm.

The main findings of this chapter are summarised below:

- The proposed system offers a high degree of flexibility since different trajectory generators, control hypotheses and arm or grasp controllers can be used.
- A grasp affordance was modelled as relationship between the robot state and a basic 3D shape, modelling an object part.
- The proposed grasp affordance definition allows emphasising on the performance of an action rather than depending on the specific object properties during grasping.
- A trajectory generator algorithm, based on models learned by demonstration, was proposed as a flexible and efficient grasp affordance controller.
- A feasibility study in a 3D simulated environment demonstrated higher flexibility and efficiency of the proposed human derived trajectory generator compared to a state of the art artificial planner.
- The trajectory generator was assessed against two control hypotheses - an approximated and a maximally reliable. The most effective and less noisy approach is a mixture of the two models.

The RRT planner has proven to be better at generalising to different actions, indeed by definition a planner must be able to reach a single point under any constrain. The main disadvantage of using a generic planner is that it does not take into account the complexity of grasping. It cannot easily alter its profile once computed, and the plan cannot be composed of the output of sub-planners: it can only reach for a specific defined point. The grasp affordance trajectory generator, instead, can be composed of different models that can be changed and adopted to the specific problem, even at runtime if needed. The closed form nature of the control models grants high efficiency and faster calculations compared to an algorithmic planner and can be computationally optimised during compilation time. The main limitation of the trajectory generator is that it relies on models tailored to a specific action. It is however believed that generalisation across actions it is easier to achieve than generalisation across objects. Indeed, some actions, such as hammering a nail and inserting a pin, are very similar to each other in terms of range of motion and assessing criteria.

Chapter 6

A Biologically Inspired Grasp Controller

6.1 Introduction

This chapter proposes a grasp controller, inspired by biological findings on human grasping described in this thesis, which can be integrated in the Grasp Affordance Framework described in Chapter 5. Such controller is designed to be object independent, in order to decouple the geometry and characteristics of the object from the fingers control, and it is not tailored to any specific hand hardware as long as it has an opposable thumb. The only discriminant factor is the initial grasping posture adopted by the robot hand and the position of the thumb, as it entails a different set of implementable grasp affordances. As demonstrated in Chapter 4, the position of the thumb in human grasping is fundamental when shaping the grip and the validity of this principle is explored here in robotics. The purpose of this chapter is to describe the grasping component of the proposed implementation of grasp affordances and to assess its performance on a state-

of-the-art robotic hand using as class of grasp affordances those that require the thumb placed in opposition to the fingers and palm.

The proposed grasping algorithm is inspired by the fact that, before clamping on the object, human digits undergo a preshaping phase [151], where an initial grasping posture is determined, and that the fingers are displacing synergistically when grasping [44], as already explained previously in Section 2.2.2. The proposed algorithm implements the preshaping and clamping of the digits as two separate stages: a preshaping stage, where fingers are displacing as indicated by a grasping synergies model, and a kinematic enveloping stage, where the digits are clamping on the object until it is firmly hold in the hand. The purpose of the second stage is to perform a grasp that restrains very different objects from the same starting point of the digits, extending the grasp generalisation of a posture in a similar way as humans do.

The original model of grasping synergies described in [44], uses a Principal Component Analysis (PCA) decomposition of the matrix containing the angular displacements (rows) for every hand joint (columns). It is proven that using a subset of synergies (eigenvectors of the matrix) creates stable grasp postures with much less control variables than a fully actuated hand [66]. Hence this model is often used as a base for a customised hand hardware design with a number of synergies fixed a priori. The use of PCA does not take into account the evolution of the synergies over time. In a different study [172], this limitation was overcame applying the Singular Value Decomposition (SVD) [173] to the human data when grasping real objects extracting eigenpostures, eigenvalues, and time modulation vectors. Such vectors are explaining how each postural synergy evolves over time offering temporal information on the role of each synergy in grasping. This model of grasping synergies is to be preferred to the original one since it describes the evolution of the synergies over time and it is better suited to a grasping controller which is dynamically executed in software and in real time.

By implementing grasping synergies at software level, it is removed the need of customizing the hand hardware and deciding and fixing beforehand the number of synergies used. It is still required to learn a finger displacement strategy from human demonstrations and to adapt and process the collected data for implementing a general purpose grasping algorithm. Grasping demonstrations can be performed in many different ways [112]. Some are intrinsic to the robot joint space - e.g. tele-operation or kinaesthetic teaching. Others rely on observing extrinsic data provided by the motion of an external body - e.g. motion capture or observational learning [174]. Furthermore, the collected data must be fitted into a suitable model in order to ensure their repeatability, such as dynamic motion primitives [163] or, in this case, grasping synergies. A similar example of learning from demonstration applied to grasping synergies can be found in [175] where authors use PCA as a model for the grasping synergies. As already mentioned, such technique does not describe well enough the evolution of the synergies over time as the SVD does. Additionally, the authors of [175] implemented the learning part using the motion capture of human actions. Since the human body is more complex than robot kinematics, there is a substantial kinematic mismatch between the body of the teacher and the one of the robot. This problem is called the *correspondence problem* [113], and it is not trivial to compensate for such discrepancy.

To circumvent the correspondence problem, kinaesthetic demonstrations are performed. Such form of learning takes place in the robot joint space and does not require any mapping between human and robot kinematics. Additionally, multiple grasping executions are required to be collected to produce a dataset suitable for learning. As grasping synergies describe the motion of one execution, several demonstrations are clustered together using K-means and a general policy is derived by interpolating the resulting centroids through splines. The robot was taught to grasp a plain cuboid since its shape is simple enough to produce a reusable set of digits' motions that can be applied to very different objects from a telephone receiver to a fencing handle.

The proposed algorithm is designed to be used on different hands with an opposable thumb and as such it was validated on a state-of-the-art robotic hand - the iCub hand. It is therefore useful to evaluate how the kinematic mismatch between the human and robotic thumbs impacts the execution of the grasp and how emulating the dominant role of the thumb when clamping the object benefits the solidity of the grip. This step is required prior including the approaching part of a grasp affordance to appreciate how the two phases complement each other.

The contributions of this chapter are the following:

1. A software grasp controller is proposed. The controller implements time modulated grasping synergies and exploits the role of the thumb in grasping to generalise on different objects. Those concepts are implemented as two separate control stages: preshaping stage and kinematic envelopment stage.

2. A general preshaping strategy is derived from a dataset of kinaesthetic grasping demonstrations by clustering the data, using K-Means, and interpolating the resulting centroids in a motion sequence using splines.

3. It is demonstrated to what extent the proposed method handles the inherent variability of the kinaesthetic demonstrations

4. It is discussed which of the two stages of the controller, taken in isolation and together, contributes most to the grasp generalisation over different objects. It is also analysed the minimum number of synergies required for a good grasp.

5. The overall performance of the algorithm and of each stage is analysed.

6.2 Learning Grasping Synergies from Demonstrations

Since the focus of this chapter is on grasping, only the 9-finger joints are considered for the analysis. Reaching and wrist orientation alignment on the robot platform is performed by the Passive Motion Paradigm algorithm (PMP) [121] at this stage.

6.2.1 Kinaesthetic Teaching Setup

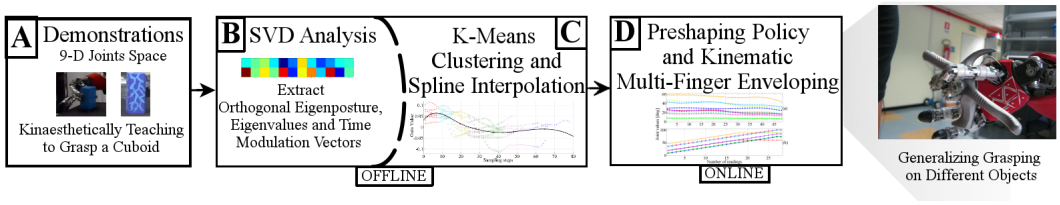


Figure 6.1: Summary of the grasping algorithm. [A] Kinaesthetic demonstrations are collected from the joints of the robot. [B] De-noised data are decomposed in orthogonal primitives, eigenvalues and time modulation vectors through SVD. [C] Time modulation vectors from different demonstrations are combined through K-Means and spline interpolation. [D] The preshaping policy resulting from [B] and [C] is executed; Kinematic Enveloping is performed.

Experimental data of kinaesthetic teaching on the iCub robot were collected from joint encoders in the arm (7 joints) and fingers (9 joints), contact sensors at the fingertips, and visual information from the two cameras as illustrated in Figure 6.1.A. Kinaesthetic teaching was performed by a human operator standing next to the robot, as shown in Figure 6.2. In each demonstration trial, the operator held the hands of the robot with the motors switched off, and guided them to perform a pick and place operation. Objects were placed within an easily reachable workspace of the robot.

The nine joints in the robot's hand were organized as follows: one joint controlling adduction/abduction of the digits, three joints for the thumb, two for the index and middle fingers and one for the ring and pinky fingers whose motion is coupled. Each demonstration trial had a minimum of 96 to a maximum of 156 data samples. In total

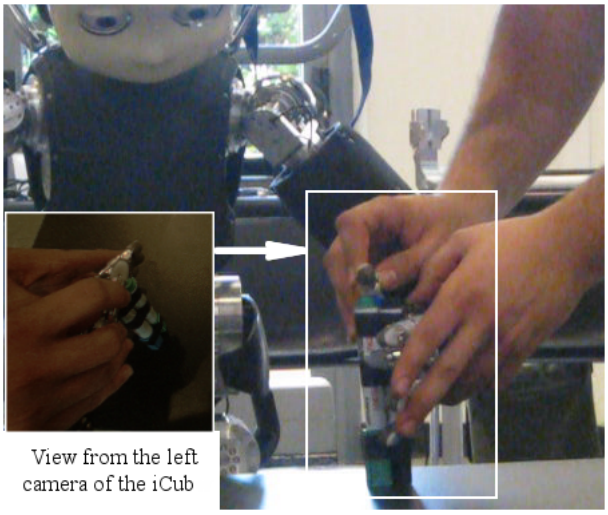


Figure 6.2: The iCub kinaesthetically learns to grasp three markers bundled together while the left camera of the robot is recording the scene.

eight trials were performed of picking and placing a cuboid as primitive object.

6.2.2 Extraction of Synergies from Demonstrations

Data Preprocessing

The data recorded from the joints of the hand have two trails of redundant stationary data when reaching towards the object and when retraction to the home position. These redundant data were automatically removed using contact initiation and termination of the tactile sensors, in order to make the postural primitives accurately represent the dynamic range of the grasping behaviour. Then it was used a 3rd order Savitzky-Golay filter [176] with frame size equal to 9 to remove high frequency artefacts in the raw joint data generating an $N \times M$ joint matrix J of preprocessed data. In this case, $N = 9$ was the number of joints in the hand, and $M \in [96, 97, 98, \cdots, 156]$ was the number of data samples in each demonstration trial.

Calculation of Synergies

This step is illustrated in Figure 6.1.B. Here, a singular value decomposition was performed on the preprocessed joint data in J in order to obtain three matrices as given by:

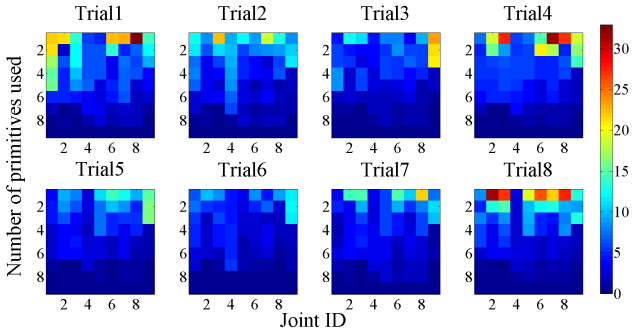
$$J = USV^T \tag{6.1}$$

Where U is the $N \times N$ matrix of the eigenpostures, S is the $N \times M$ matrix of the eigenvalues, and V is the $M \times M$ matrix of the temporal modulation coefficients for each primitive in U. In this context, matrix U represents N orthogonal postural primitives of the hand that span the variability of the corresponding grasping demonstration. Each element in the diagonal matrix S represents the significance of the corresponding postural primitive. The higher the values, the more the eigenposture contributes to the variability of the digits. The columns v_i in matrix V represent the scale of temporal modulation of corresponding primitives in U.

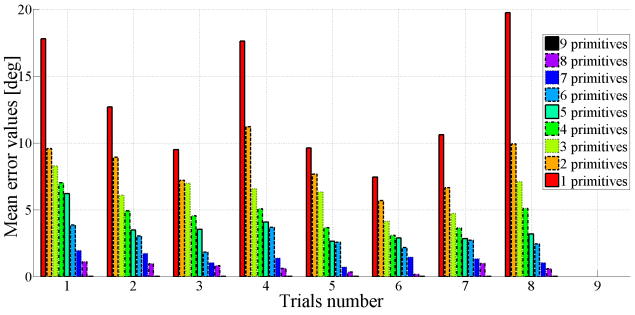
In order to decide how many primitives to consider in the analysis, the Root Mean Squared Error (RMSE) is evaluated. The RMSE for each trial i is calculated for $k = 1, 2, \cdots, 9$ primitives as given by:

$$RMSE_i^{(k)} = \sqrt{\frac{[(U^{(k)}S^{(k)}V^{(k)^T}) - J]^2}{M}}\,, \tag{6.2}$$

Where $U^{(k)}$, $S^{(k)}$ and $V^{(k)}$ are the posture primitive matrix, the significance matrix, and the time modulation matrix respectively, for k number of primitives, J is the matrix of preprocessed joint data profiles, and M is the number of samples recorded in trial i.



(a)



(b)

Figure 6.3: Figure 6.3(a) are RMS errors per trial. Each column shows the error on a single joint, while each row shows the error using a different number of primitives (from 1 to 9). Tables have been normalized to 35% to improve visualization. Figure 6.3(b) shows the distribution of mean RMS errors using a different number of primitives.

Figure 6.3(a) shows the distribution of RMS errors over the trials when varying the number of primitives for each joint. The values have been normalized to the maximum of all trials. By observing the errors, it can be concluded that a good approximation could be obtained by considering two or three primitives only. The maximum error among all trials is 11.6% for two primitives and 9.5% for three primitives (Figure 6.3(b)). Those results are in line with the observations of Santello et al. [177].

A comparison between the reconstructed joints and the de-noised data can be observed in Figure 6.4(a) and Figure 6.4(b) for two and three synergies employed respectively. It can be noted that employing more synergies reconstructs a preshaping policy more

similar to the original data, especially for some joints.

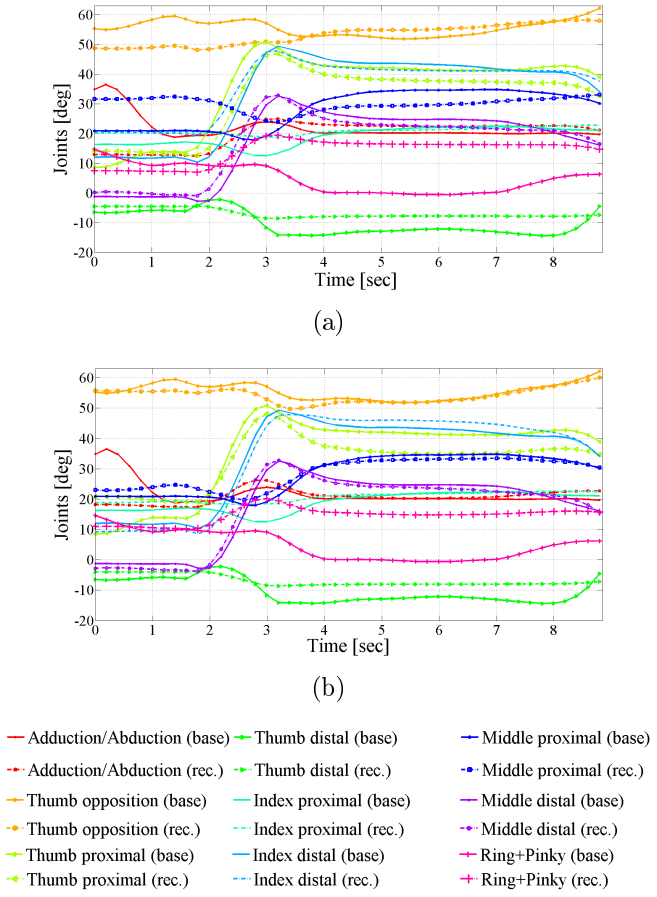


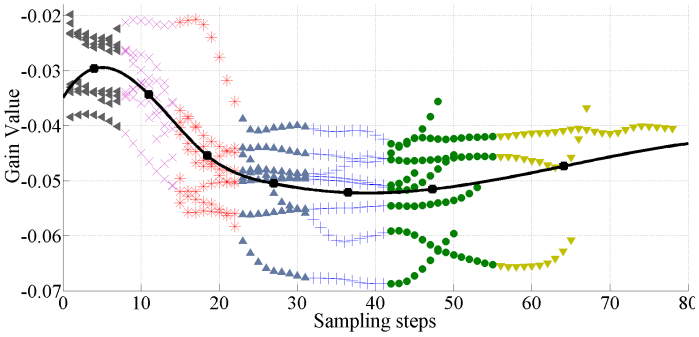
Figure 6.4: Comparison between de-noised joint values of trial 6 (continuous line) and reconstructed ones (dashed line) using **(a)** two primitives and **(b)** three primitives.

Integration of Different Teaching Trials

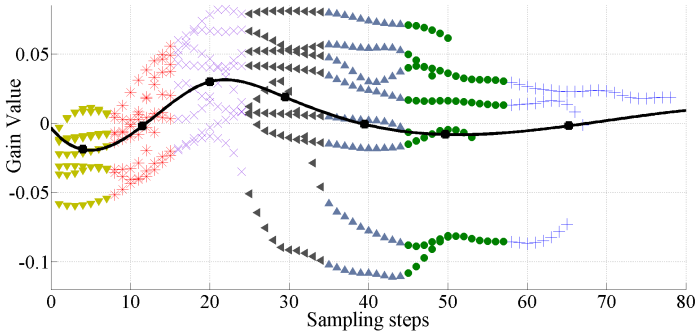
This step is illustrated in Figure 6.1.C. Since eight demonstration trials have been performed for grasping the same object, the problem of abstracting them to one set of postural primitives and corresponding time modulation vectors has to be solved. The pre-processed joint data matrices of each trial were combined into one single matrix, and SVD decomposition was applied on it. Concatenating the joint matrices of different tri-

als poses the problem of separating the contribution of each individual sequence when extracting a unique set of time modulation vectors. This problem of abstraction was solved by performing K-Means clustering [178] of the ensemble of $v_i, \quad i = 1, 2, 3, \cdots, 8$ vectors from the 8 demonstrations as shown in Figure 6.5. Centroids of clusters (*milestones*) are used as key points to build interpolating cubic splines to reconstruct an abstracted preshaping policy. The longest demonstration was used as a reference for determining the number of clusters ($K = 7$) of the K-means clustering algorithm. Each eigenposture has an associated best fit set of splines that encodes a preshaping policy.

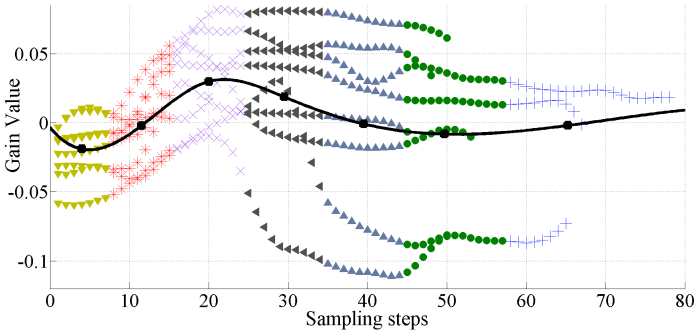
Thanks to the orthogonality property of SVD, each component of the v_i matrices can be interpolated separately. It is therefore possible to simplify a 9-dimensional control algorithm to two or three independent spline interpolation problems.



(a)



(b)



(c)

Figure 6.5: Results of K-Means clustering on temporal modulation vectors. Vectors v_i 6.5(a) $i = 1$, 6.5(b) $i = 2$, 6.5(c) $i = 3$ from different trials are clustered independently in order to obtain a set of 7 key points for spline interpolation. Cubic Spline interpolation from resulting milestones is shown as black curve for each dimension.

6.2.3 Kinematic Multi-fingered Enveloping

The final stage of grasping is the kinematic enveloping. Postural primitives extracted from the SVD analysis and the corresponding spline interpolations of the temporal modulation coefficients were used to compute the final preshaping grasp policy realtime. This step is summarized in Figure 6.1.D.

Once the hand is preshaped by following the joint values encoded in the grasp policy, each finger wraps around the object trying to squeeze it. Joint encoder readings collected over a given time horizon ($t_f = 10$ sec) were used to detect the success of a grasping sequence. This enhances the portability of the algorithm to different robotic platforms that may or may not have tactile feedback on the fingertips.

A digit is considered to be steady if:

$$d_i \in [f_i(l)|h_i < x_f], \tag{6.3}$$

where d_i is the total angular displacement of digit i, f_i is the function regulating the motion of joint i, k is the current number of encoder data samples, h_i is the total angular displacement of the i-th joint in a given time horizon ($t_f = 10$ sec), x_f is the threshold total angular displacement to detect whether a significant angular displacement has occurred in each time horizon.

The execution flow is resumed in Algorithm 2. $c_i = 2.5°$ is the amount of angular increment of i-th joint in each sampling step $T = 1.6$ sec. The increment c_i has been selected from the mechanical constraints of the robot to ensure an observable motion due to the resolution of the encoders. The time horizon t_s to detect a terminal condition

Algorithm 2: Kinematic Enveloping Algorithm

Data:

t - execution time $[0 \cdots \infty)$ sec
h_i - total displacement of joint i
f_i - motion of joint i
l - number of iterations of the algorithm
x_f - threshold defining if the motion of a joint is negligible
c_i - constant increment of joint i
T - sampling time
t_f - temporal horizon when stopping criteria is evaluated
P - number of joints whose motion has been detected as negligible
$Jlim_i$ - limit for joint i

```
begin
    P = 0;
    for l = ⌈t/T⌉; f_i < Jlim_i do
        for i = [1 ··· J_n] do
            f_i(l) = l c_i;
            h_i(l) = Σ_{q=0}^{q=l} f_i(q);
            if l = ⌈t_f/T⌉ and h_i(l) < x_f then
                P = P + 1;
            end
            if P >= 6 then
                return true // Grasp successful
            end
        end
    end
    return false // Grasp unsuccessful
end
```

is set to 10 seconds, while x_f is set to 10°. The time horizon t_f and the threshold x_f are the functions of the increment and have been derived from tests on different objects. If the number of stationary joints is $P >= 6$, the object is considered grasped. This number has been evaluated experimentally, based on the assumption that at least three digits are needed to grasp tightly [179] and that about two joints are controlling a single finger. The quality of the grip is validated experimentally as illustrated in next section.

6.3 Experimental Evaluation

In this section, the proposed algorithm is validated on a state-of-the-art robotic hand by grasping real-world objects. At first the robotic platform, iCub, is introduced. Secondly, the experimental set-up and the validation criteria are described. Finally, the results of the algorithm are discussed.

6.3.1 Description of Robotic Platform

The iCub is a 104 cm tall humanoid robot resembling a 3.5 year old child [180] with 53 degrees of freedom in total - including seven for each arm and nine for each hand. The hands are actuated with a tendon driven system with the motors located in the forearm (Faulhaber 1016M012G). Encoders are based on hall sensors. The hands of the robot are equipped with capacitive pressure sensors [181] that are distributed on palms, fingertips and the rest of the body. Twelve taxels are allocated on the fingertips, and detect the contact with conductive material (such as metal or human skin).

The robot, apart from the Application Program Interface that communicates directly to the hardware, is operated through YARP [182], an open-source robotic middleware that supports distributed computation.

6.3.2 Experimental Setting

In order to validate the algorithm, several experiments have been carried out with the iCub grasping different objects. The objects used in the experiments are shown Table 6.2. The hand was positioned at a fixed maximum distance of 5 cm from the grasping point in order to manage a successful grasp.

Afterwards, the algorithm is executed and the preshaping is performed. Then the kinematic enveloping phase takes place. This technique differs from other methods [183] based on known contact points, therefore the concepts of force/form closure [97] are not appropriate methods to validate the stability of the grip as the position of the contact points is not known a priori. Instead, validation has been carried out experimentally. Once the algorithm terminates, the object is vertically lifted at about $22 \cdot 10^{-3} m/s^2$ and a human operator pulls the object from the robotic hand to verify the stability of the grip (Table 6.1). Each grip has been ranked based on the force exerted by the human operator, it has been estimated experimentally using a 6 DoF force and torque sensor (MINI40, ATI Technologies).

Index	Description	Median of Force (N)	Standard Deviation of Force (N)
4	The object is removed with a significant effort from the operator, possibly using two hands and manually opening the robot's hand.	92.35	7.23
3	The object is removed with significant effort from the operator but without interacting with the robot's hand.	51.59	5.88
2	The object is removed sliding it from the robot's hand applying discrete effort with no interaction with the robot's hand.	15.81	1.48
1	The object is easily removed with little effort and without interacting with the robot's hand.	4.9	1.86
0	The object is not grasped or is falling while being picked up.	-	-

Table 6.1: Description of the ranking score for grip validation.

6.3.3 Experimental Validation and Discussion

In order to validate the robustness and generalization power of the proposed grasping technique, all objects in the set are grasped with different numbers of synergies: three, two, one or just with the kinematic enveloping.

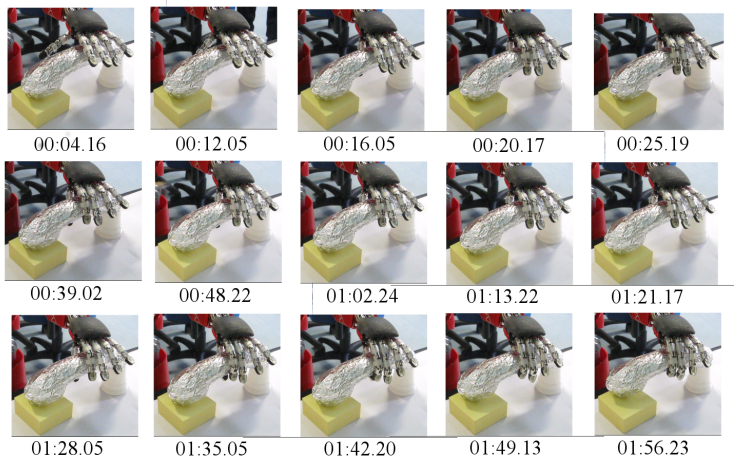


Figure 6.6: Evolution of the grasp on the real robot using two primitives for preshaping. The time reported in the first picture takes into account the required computation to start the algorithm.

To analyse whether different demonstrations are able to produce a feasible preshaping policy, despite the variability of the teaching trials, it is possible to observe the evolution of a grasping sequence using just two primitives as in Figure 6.6. It can be observed that the merge of different teaching demonstrations executed on the same object - the cuboid - produced a stable grasping policy. Thanks to the K-Means clustering and the spline interpolation, it is possible to combine different demonstrations in order to produce a unique preshaping policy.

Figure 6.7 shows the reconstructed evolution of the joint angles during the grasp execution, while grasping a telephone receiver using three synergies. As explained in section 6.2.2, the preshaping follows a policy defined from the extraction of synergies from kinaesthetic data. Therefore its evolution is very similar across different objects. The

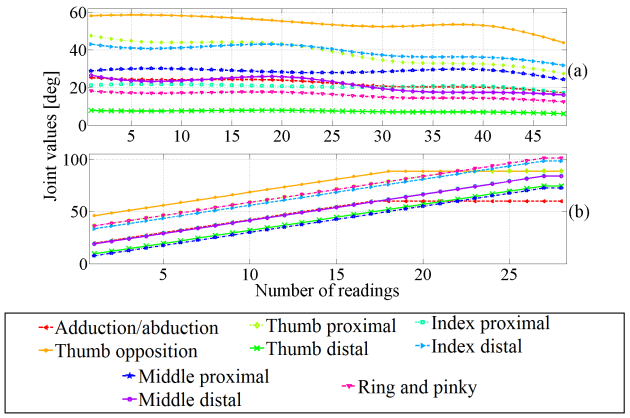


Figure 6.7: Evolution of joint values for grasping a telephone receiver using three synergies. **(a)** Evolution of the joints of the hand during the preshape phase. **(b)** Evolution of the joints of the hand during the kinematic enveloping phase. The fingers are closing linearly until the object's surface is hit. Afterwards, the motion reaches a steady state and the grasp termination is detected.

enveloping phase, however, differs from object to object. It can be clearly seen that, as defined in section 6.2.3, the joints are linearly closing until the object is enveloped with the fingers reaching a steady state.

Table 6.2 summarises the results of grasping the objects. It can be seen that the telephone receiver, the markers, the cylinder, the cuboid and even a complex shaped fencing handle, are tightly grasped (see grade of stability in the third column of Table 6.2 and Table 6.1 for the description of grades of stability). The reason is that their geometries allows the use of the same grasp configuration as the object used for learning and the thumb is placed on the right position - ranked 4 to 3 on Table 6.2. The computer mouse is grasped but the grip is not tight - ranked 2 on Table 6.2.

Object Name	No. of Synergies	Grade of stability	Duration [sec]	Preshaping time [sec]	Enveloping time [sec]
 Cuboid*	3	4	78	35	43
	2	3	92	37	55
 Cylinder*	3	4	133	30	103
	2	4	89	26	63
	1	3	111	30	81
	Envl. Only	2	119	-	119
 Receiver	3	4	109	73	36
	2	4	109	74	35
	1	2	109	74	35
	Envl. Only	2	56	-	56
 CD Case	3	1	124	74	50
	2	1	120	74	46
	1	1	127	72	55
	Envl. Only	1	48	-	48
 Markers	3	4	119	73	46
	2	3	119	74	45
 Mouse	3	2	116	77	39
	2	2	109	74	35
	1	1	109	74	35
	Envl. Only	0	4	-	4
 Glass	3	0	111	73	38
 Handle	3	3	122	77	45
	Envl. Only	3	59	-	59

Table 6.2: Summary of grasping experiments on the set of objects. (*) The time has been evaluated using different parameters for the kinematic enveloping phase ($c_i = 1.0$, $t_c = 0.5$), moreover the termination of the enveloping was manual.

The Compact Disc (CD) keep case was not successful since its geometry suits better a parallel grasp [38]. This posture requires a precise contact timing of the thumb and the other fingers and the enveloping phase was not complying to it as it is a linear motion. Alternatively, the position of the palm could have been adjusted to guarantee the right timing for a linear motion. As the shape of the glass is cylindrical, it was more interesting to test a grasp with splayed fingers on the top of the glass. However, the grasp was unsuccessful as it was difficult to place the thumb and the other digits in such very specific configuration due to kinematic limitations.

Based on this experimental studies, it was found out that a misplacement of the yaw angle of the opposition joint of the thumb (angle between the thumb and the index finger) is the cause for most of failures. This conclusion is logical, because this degree of freedom of the thumb has given primates the skill to manipulate objects [184] and it is in agreement with the related findings in this thesis. Other digits played a more supportive role in achieving a good grasp.

The experimental results given in Table 6.2 confirm that the difference between three and two postural primitives is negligible. Both design choices perform comparatively well, as the thumb opposition placement is similar. Using only one synergy causes failures due to the very large dimensionality reduction and its consequent reconstruction error. The thumb is often misplaced leading to an unsuccessful grasp.

The enveloping phase has been proved to be very helpful but not sufficient to finalize the grasp. Unless the object geometry is very simple, enveloping only may cause failures due to the blind placement of the thumb on the object. However, with a proper initial configuration of digits after the preshaping phase, the enveloping helped to generalize for cases where the size of the object was smaller or the geometry was different from the original object used to demonstrate grasping.

Among the two parts of the algorithm - preshaping and enveloping - it can be seen that the second stage is generally much faster than the first. The reason is that the fingers move quickly and generally the enveloping is executed after preshaping, so the number of iterations cannot be very large. The preshaping phase, instead, is longer since it depends on the number of sensor readings that, in general, can be very large depending on the demonstration.

In the case of the cuboid, the overall time is limited, since the list of joints commands is short (about 48). Indeed, due to mechanical constraints on the robot's hand, a brief pause is required between commands (1.5 seconds in the experiments) to enforce the execution of the command. This problem can be easily circumvented by sub-sampling the policy observing the gradient of each joint. If the derivative is not large enough, it can be deduced that the requested motion is too fine to be achieved and that set of joint commands can be skipped in favour of larger values.

6.4 Conclusions

This chapter proposes a biologically inspired algorithm for grasping based on kinaesthetically learned grasping synergies and on passive kinematic envelopment of the target object. Although the dataset was based on demonstrations of grasping a simple cuboid, the robot was able to grasp well objects very different from the original, such as a fencing handle and a telephone receiver. This was possible because the kinematic envelopment adapted the original posture of the synergies to each different object, assuming that the thumb was correctly positioned for the required grasp.

Experimental results show that enveloping in itself is not enough to ensure a safe grip in all cases due to the blind placement of the thumb opposition joint. Therefore, in the general case, it is required to *bootstrap* the thumb to an initial position, which possibly

can clearly differentiate a Key Grasp from a Power Grasp.

The robotic experimentations confirmed the dominant role of the thumb in grasping. Unsuccessful grasps were often caused by a misplaced thumb or an incorrect contact timing between the thumb and the other fingers. On the other hand, the fencing handle was grasped very tightly although its geometry is very different from a cuboid. Again this was possible because the thumb was placed correctly. It can now be seen that an inadequate rendering of the human kinematics of the thumb in robotics limits sensibly the scalability of a given initial posture when enveloping different objects. A more dexterous thumb might have allowed grasping some difficult objects or improving the grip of others.

The proposed implementation of the controller is in line with its biological counterpart also because the preshaping phase can be executed in parallel with the approaching motion, optimising the execution time. Finally, it can be seen that the grasping component of a grasp affordance does require an adequate alignment to the object in order to perform correctly and reduce the failures. The palm alignment can be especially helpful to minimise the fails related to wrong contact timing of the digits, as the hand can be closer to the object to reduce the contact time of the index and middle fingers. Humans can be taken as base model given their advanced manipulation skills. Hence it is clear that the contribution of a biologically motivated approaching strategy is a mandatory component for the implementation of grasp affordances in state-of-the-art robotic hardware.

The main contributions of the chapter are listed below:

1. The 9-dimensional joint space of the hand of the iCub robot can be reduced to two or three postural primitives that can be independently combined. Experiments have shown that the choice between two or three postural primitives does not

have a considerable impact on the final quality of the grasp.

2. The use of kinaesthetic teaching as learning technique intrinsically solves the correspondence problem, and saves from the burden of mapping the human kinematics as in other approaches [66], [175].

3. Data collected from different teaching trials can be integrated through the use of K-Means clustering and spline interpolation, regardless of their variability. The final result is a unique set of time modulated synergies used for preshaping the hand to an initial posture.

4. From an initial preshaping configuration, the fingers are wrapped around the object until no significant motion is detected. This phase, called kinematic envelopment, allows to achieve a more general grip that is able to scale on the target size and geometry of the object.

5. The kinematic limitations of robotic thumbs require to be compensated with a biologically motivated approaching strategy that would enforce an appropriate timing and positioning of the thumb on the surface of the object.

Chapter 7

Conclusions

This conclusive chapter provides an overview of the thesis. The main contributions are listed and summarised in Section 7.1, and potential industrial applications of the findings and the proposed method are presented in Section 7.2. Additionally, in Section 7.3, the main assumptions and limitations of this thesis are listed and discussed in the form of requirements. Moreover, proposals for other aspects of grasp affordances not covered in this thesis, such as learning and perception, are illustrated in Sections 7.3.1 and 7.3.2.

7.1 Summary of Contributions

This thesis contributes to the field of multifingered grasping for robotics, setting the base of a novel approach to grasp affordances that relies on action dependent requirements rather than on specific object features. The main motivation of this approach is the need for robotic grasping and manipulation of very different objects, without the complete re-calibration of a complex grasping system or learning control strategies for every novel object. The state-of-art approaches are often domain specific and require

repeated learning or complete re-design of a synergic robotic hand. The proposed approach is justified by novel and state-of-the-art results in human neuroscience of motor control. Studies of human behaviours for grasping and approaching are presented and studied to outline the cardinal principles of the presented architecture and controllers. This thesis is a first step towards a more structured and modular organisation of a grasping system that is required to enhance grasping effectiveness. Such architecture separates the problem of motor control from the challenges of perception and learning, so that each component can be explored and developed and combined independently. The overall validity of the proposed approach is evaluated for robotic implementations, both for simulation and real setup, demonstrating and taking into account the grasping limitations of the state of the art robotic hands.

More specifically, the thesis provides the following contributions:

1. In Chapter 2 a thorough introduction to human neuroscience of motor control, aimed at a robotics audience, is given and related to the state of the art research in robotic grasping and affordances. Limitations of state of the art approaches are highlighted.

2. In Chapter 3 human behaviours of approaching to grasp to hammer are studied and characterised in terms of hand-to-object distance and orientation. Different models of approaching, for a robotic implementation, are derived and validated.

3. In Chapter 4 the dominant role of the thumb in human grasping is established by statistical analysis of human grasping data. The limitations of state of the art hands, in the context of thumb dexterity are demonstrated, and suggestions for improvements are given.

4. In Chapter 5 a modular architecture for grasp affordances is presented. Criteria to model successful grasp affordance strategies are outlined, and a grasp affordance

trajectory generation algorithm, derived from human studies of approaching to grasp, is provided and evaluated in a feasibility study during simulations.

5. In Chapter 6 the findings on the role of the thumb are applied to a software implementation of a synergy-based grasping controller on a real robot. It is demonstrated that a correct placement and timing of the thumb allows the controller to scale over different objects and how such controller fits in the general architecture.

7.2 Practical Applications

In this thesis, the formulation and implementation of grasp affordances as a modular architecture are envisaged as a tool for domain-independent grasping that can be applied to vast range of different fields. Given the generalisability of the potential applications, some practical aspects for concrete industrial settings are described in this section.

One of the fields of primary importance is manufacturing and industrial applications. Some industries, like automotive or die-casting[1], already rely on robotics and it might be easier to integrate the results proposed in this thesis, building up on the existing knowledge and infrastructure of those industries. Additionally, a robot able to grasp and flexibly manipulate a large range of objects can be introduced in a Just-In-Time (or flexible) assembly line, diffused among many manufacturing businesses[2], where workers interact with a large number of components to assemble a customer-tailored product. A set of assembly actions can be learned and composed in an action plan, and the proposed system can perform the same operation using different tools, or it could grasp the same tool in different ways.

[1] Buhler Druckguss AG, Uzwil, Switzerland
[2] Buhler UK Ltd, London, UK; Toyota Motor Corporation, Aichi, Japan

A robot giving the human a supervisory and monitoring role can perform repetitive actions in an unstructured environment, where the unpredictability of the position and availability of tools is a challenge. Such system might reduce the stress and inefficiencies and increase the ergonomics for human workers. Those are challenges for many manufacturing techniques [185]. Without a flexible manipulation system it is required to know the position of every object in the workspace, as it is dictated in traditional industrial robotics. Logistic automation industry can also benefit from the proposed approach, as an object independent grasping system can be used to pack goods[3] or groceries[4], setting free the human from the burden of manual packing. In the same way as commercial needs were driving the evolution of computer science and networks in the 1970s, industry and manufacturing can drive the evolution of robotics in future. Therefore, the solutions to the problem of grasping and affordances should satisfy the requirements of flexibility, performance and cost for each different application.

Another domain of use for the results of this thesis is manipulation in extreme and dangerous environments, such as disaster response, space exploration or nuclear decommissioning. In the Darpa Robotic Challenge (DRC) [186], for instance, a robot was required to navigate in a dangerous environment and to interact with numerous objects and tools, such as hammers, doors and valves, in order to perform tasks normally performed by human rescue teams in the field. The proposed system could enable the robotic rescuer to manipulate novel objects segmented in grasping areas associated to known grasp affordances. A similar remark is valid for space exploration missions, such as NASA's Robonaut project [156], where the robot operates on International Space Station. Another example is nuclear decommissioning[5] , where the choice of objects is more defined, but the hazards and the required manipulation dexterity do not allow traditional robotic solutions or human intervention . The additional challenge of the

[3]Amazon Robotics LLC, Massachusetts, USA
[4]Ocado Technology LSE, Hertfordshire, UK
[5]Oxford Technologies, part of Veolia Group, Abingdon, UK

above applications is the prohibitive or life threatening environment where the robots operate. Extreme heat or cold, unstable surfaces, lack of gravity, ionising radiation and other characteristics make the environment dangerous and unsuitable for a human operator. Flexible manipulation capabilities can be an excellent feature for such robotic systems. However, better engineering solutions are needed to guarantee the required high execution performance and reliability.

Another possible domain of application of this work is service robotics. Robots are required to operate in unstructured environment, such as a household or a nursery, where no prior assumptions on the surroundings can be made. The potential for application of the proposed results is very wide: from tidying up the environment to preparing a meal or cleaning the house. The wide range of possible actions and objects in different configurations and orientations make such environment a challenging opportunity for benchmarking and extending the proposed architecture. Despite the great potential of applicability of this method, it is advised to use such domains as test scenarios for novel scientific results and extensions of the present system. The reasons for this are of technical, legal and social nature. On the technical point of view, a robot that requires to be frequently calibrated or to be taught many different actions is very uncomfortable to use, if it is not supported by adequate scientific research and engineering solutions. In addition, such system might fail, as it was the case for the first commercial speech recognition software in late 1980s [187]. On the legal and social point of view, the lack of regulations and understanding from the general public can pose a threat to the actual human user. It is yet unclear how to define liability if a robotic system causes an accident due to contingent factors, such as an unexpected lack of sensory information. Standard and universally agreed default behaviours have yet to be defined as a mean to mitigate the effect of uncertainty and to promote safe human-robot interaction. If it is known a priori the fail-safe behaviour of a robot, it is easier to prevent trivial accidents such as a child's hand being accidentally squeezed when interacting with a

sensory-blinded robotic hand. Common knowledge of a robotic system should be first established in society to avoid the occurrence trivial hazards as described above. In the same way it is obvious that a child shall never be allowed inserting the own fingers in the mains socket. All the mentioned points are much more difficult to settle than technological advances. The risks and hazards derived from a rushed introduction of robotics to the general public might reduce the users' confidence in robotics and even create unjustified concerns.

It is therefore recommended to apply the results of this thesis to industrial and manufacturing problems initially. Robotic applications in extreme and dangerous environments can be used as an opportunity to better engineer the proposed framework and required technologies and to pioneer uses in less structured environments . It is point of view of the author that applications in service robotics shall be approached as an open research problem, to expand new horizons and to explore completely new routes. In this regard, the results of this thesis can be an initial attempt to structure the problem of grasp affordances, which can be further expanded with additional research.

7.3 Assumptions and Future Research Directions

This thesis is a first step towards a more general purpose grasping system, and a more structured definition of grasp affordance. Therefore, some assumptions were made and some requirements should be fulfilled in order to transfer and to use the results for real applications. Such constraints can be considered as opportunities for future improvements. As the roles of perception and learning were not investigated in detail, and a more comprehensive discussion is required. For this reason, the motivations for this choice and proposals for extensions are discussed in Sections 7.3.1 and 7.3.2 of present chapter.

Requirements and Assumptions

The proposed architecture of a grasp affordance system was derived from studies of human motor control and assessed on humanoid robots using everyday objects. As such, the following requirements should be fulfilled for this approach to be easily transferable with little effort:

- A humanoid multifingered hand should be used;
- The volume of the reachability space of the robot should to be comparable to the human counterpart;
- A template action should be based on a motion taught from human demonstrations;
- Hard, human-made objects should be manipulated;
- The action shall be performed in absence of obstacles;
- The result of the performed action must be assessed in closed form;
- Complex actions, such as object assembly, should be broken down in simple assessable actions;
- A graspable object should have at least one grasping area represented as a primitive 3D shape;

The above requirements are providing guidance that can enable easy transfer of the results of this dissertation to other similar platforms and domains. The above conditions might be relaxed or extended if a different robot or domain is considered, preserving the validity of the approach. For instance, if a suction-based gripper is employed, the synergistic grasping controller can be substituted with a simpler one. On the other

hand, handling deformable objects might require a more advanced multifingered grasping controller that takes into account the stiffness of the target. In both examples, the rest of the system, such as the arm controller and the trajectory generator, might not require any substantial change, thanks to the modularity of the system.

It is also important to underline that some assumptions and constraints were made in this work. One of the main assumptions is that most of multifingered grasp postures require the use of the thumb but some, classified as Primate Grasp postures in Chapter 4, rely only on the other four fingers. Such postures were not considered since they are not often used in humans and few everyday objects might be grasped only in that way. Therefore, as the proposed grasp controller relies on the use of the thumb, it is unlikely that it can handle such postures effectively.

A second observation is that the motion perturbations introduced by obstacles and perceptual errors were assumed as not present. One reason is that the problem of obstacle avoidance in planning taken in isolation is well studied and understood in literature. Hence, it might be easier to include the latest results in a general grasp affordance framework, which, instead, has to be established . Obstacle avoidance is fundamental in contexts, such as cluttered supermarket shelf. However, there are various well-established techniques that can be easily integrated [119].

Additionally, it is important to underline that hard and still objects were mostly used in this thesis. It is likely that, for soft and deformable objects, an extension to the grasping controller, or a stiffness classifier [84], is required to handle the additional complexity. Since soft objects, such as food, are broadly used in daily tasks, additional research is required to manipulate them and the proposed architecture will require to be extended accordingly. Similarly, if the target object is moving, such as a mechanical part on a conveyor belt, additional components should be included to the framework. For instance, the system might require a tracking component to follow the grasping

areas, and the grasp affordance relationship might need to be expressed in the base frame.

7.3.1 The Problem of Perception in Affordance Selection

The role played by the sensors and the processing of perceptual information is fundamental in robotic systems. A robot that is unable to detect positions, poses and features of the elements in the surrounding environment cannot interact with it. Despite the primary importance of perception for biological organisms and autonomous artificial systems, both entities are struggling to unambiguously perceive the surrounding environment, often failing to act appropriately. For example, the approaching of an earthquake can be anticipated by detecting seismic waves. Such waves propagate in the ultrasonic spectrum that is inaudible by the human ear, but it is perceivable by the canine auditory system [188]. Hence, a human might not be able to react as quickly as a dog in such critical situation. Dually, the sensitivity of the dog's eyes encompasses a narrower wavelength than those of human's [189]. In addition colours such as red and green [190] are excluded from dog's vision. Hence, a dog might not be able to distinguish between a green signal and a red one at a pedestrian traffic light and use this information to avoid to wrongly cross the road at a dangerous moment . If one should consider only perceptual stimuli, the dog and the human would live in two separate worlds. This is the problem of perception.

The impossibility to discriminate the objective reality from a subjective opinion on the surroundings is well known and it has fascinated numerous philosophers. In Europe, the Greek Parmenides of Elea [191] (born 515 B.C.) first formulated this problem, which was also studied in other non-European philosophical schools [192]. He stated that the existence (being), as objective external facts, and the negation of existence (not-being) are well separated entities which are difficult to discriminate due to the limitations of

human perception, initiating a debate among European philosophers on this topic. Later on, the French mathematician Rene Descartes [193] (born 1596, died 1650), proposed a fundamental contribution to the discussion stating that only the conscious thinking of oneself is guaranteed as a certain and concrete sign of existence. No assumption can be made on any other external stimuli, which might be the result of a false positive, a misclassification or not exists at all, since the partiality of perceptual information might be a leading confusing element in one's mind. This problem is still debated and unsolved today [194] and an agreement was not found yet. It can be understood that the problem of perception is of formidable complexity, and it might sound naive to hope to find a general and comprehensive solution to an issue debated for centuries, and yet not fully defined, within a single thesis. Indeed, it is not possible to define how a sensor can unambiguously perceive the properties of an object as simple as a glass and which properties are important.

Despite all this, it is still possible to find a solution to an approximate or narrower version of the original problem. For example, the weight of a glass is more relevant than its colour in grasping, but this might not be true if the task is finding the object on a cluttered desk. For this reason, a modular architecture is a good initial approach for solving a restricted problem of perception, applied to a defined domain. Modularity allows refining or limiting the scope of the problem or the range and precision of the perception sources, while the research and evolution of other components of an affordance system, like motor control, can progress with little influence. For instance, a search and rescue drone might benefit more if the range of perceptions of a dog are privileged over those of a human. An olfactory sensor might be more useful than an expensive hyperspectral camera if the task is earthquake disaster response. Perceiving all waves up to the ultrasonics might be a more appropriate level of precision if it is required to perform earthquake prevention. However, the same precision might not be needed for a search and rescue operation, as the human voice is always within the audible spectrum.

This approach might also alleviate the problem of cost and limited number of sensors on a robot, if compared to a biological organism, as the required input is dimensioned for the task at hand. The decisions on the sensor used or the data processing methodology selected for learning and detection will not influence the rest of the system. It can be assumed that every other component requires a final and processed cognitive abstract representation of the environment. Therefore, a knowledge base, similar to the human thalamus, can decouple motor control from perception, and information irrelevant to the task can be eliminated.

To summarise, solving the problem of perception in the context of affordances is a challenging task. The intrinsic ambiguity and relativity on how the external world can be perceived, even by biological organisms, cannot be fully addressed in this thesis only. However, decoupling the perceptual means from the representation of the environment with a knowledge base, is an initial step towards a more configurable system. This would enable the exploration of different types of sensory options, offered by the current technology, and means of integrating different sensory data to provide the minimal information for a motor actio n. In this way, it is possible to narrow down the problem of perception to a specific application until better technology would be able to provide multi-modal information that enables execution of more sophisticated operations.

7.3.2 The Contribution of Learning

The affordance architecture and controllers described in this thesis build up their foundations from human learning. Training data collected extrinsically and intrinsically from a human teacher were processed to obtain the initial grasping synergies and approach models . However, the system is not able to cope well if a different requirements are imposed, such as an action comparable to hammering, or an alternative grasp posture with a different thumb placement. In its current state, the system misses the ability to

adapt. Typically algorithms are made adaptable by adjusting their parameters during execution, using techniques such as reinforcement learning, adaptive control and similar. However, such approaches are impractical for a real robot, since a large amount of training data is required from numerous trials. Performing so many trials directly on the robot is very tedious and time consuming . For this reason, often the teaching data is generated in simulation [106] and the resulting models are adapted to the robot with additional validation testing. The risk of this methodology is to obtain a performance lower than expected after the validation, since any physical environment is more complex and variable than a simulated one.

Adaptability can be introduced in the system by pursuing an hybrid learning approach, which has been first inspired from the learning approach proposed by the European philosopher Socrates (born 470 BC, died 399 BC). He suggested that knowledge is innate in every human and the task of the teacher is to extract it through discussion and promoting the student's freedom of experimentation in a process called maieutic [195]. In a similar way, the demonstrations of a teacher are processed to produce a motor control policy, such as for the approaching models or the grasping synergies discussed in this thesis. This represents the initial knowledge of the robot on the task to be performed, which is innate in the Socratic learning method. If an action similar to the one initially learnt has to be executed, an algorithm of the reinforcement learning class can derive a new control policy from the initial knowledge learnt by demonstration. This step can be equated to Socrates' maieutic learning process. The assumption is that, if the new action is similar enough to the original one, reinforcement learning shall settle down from the initial control policy to the new one faster than learning from scratch or using a starting point defined from a theoretical criteria. This methodology would further shift the attention from the specific action performed, as described in this work, to a definition of similarity between actions in the space of parameters of the model improved by reinforcement. If two actions can be ranked based on a similarity

metric, it is not fundamental to learn a specific action as this can be adjusted as needed. For instance, if hammering is the taught action, pressing, clicking and insertion are similar actions, since they all require a comparable indenting motor sequence but with different timings or applied forces. A suitable motor sequence for a similar action can be established externally, as a human supervisor confirms the validity of the grasp posture and motor planning when satisfied. The robot itself can assess the quality of the grasp affordance executed by evaluating the results produced by the affordance using a performance metric designed for the new action. If a comparison can be made, both cases correspond to the intervention of the teacher in completing the learning process according to the Socratic method.

To summarise, one limitation of the proposed grasp affordance architecture in its current state is the difficulty of generalising across different actions. Reinforcement learning and adaptive control algorithms can flexibly adapt the system to new actions. However, the gargantuan level of redundancy introduced by multifingered hands and anthropomorphic arms turn this methodology into a challenging operation often executed in simulation and later validated in real life with uncertain results. As such, initialising a reinforcement learning session from an initial human demonstration can be a viable alternative to scale and adapt a pre-existent control model to a different, but similar, action. Defining a similarity metric in the model parameter space is a real barrier to overcome, but it is believed that it is easier to define a metric rather than to learn a completely new model or a few weeks long simulated learning session. A new suitable control policy can be established either by a human supervisor or by fulfilling performance criteria. In addition, it is unrealistic to find the best possible grasp posture of an object for a robot as much as there cannot be an agreement on which one among many grasping postures of a pen is the best, if the task is to sign a legally binding document.

Bibliography

[1] T. Wynn, "Piaget, stone tools and the evolution of human intelligence," *World Archaeology*, vol. 17, no. 1, pp. 32–43, 1985. PMID: 16470990.

[2] S. H. Ambrose, "Paleolithic technology and human evolution," *Science*, vol. 291, no. 5509, pp. 1748–1753, 2001.

[3] T. Okada, "Computer Control of Multijointed Finger System for Precise Object-Handling," *IEEE Transactions on Systems Man and Cybernetics*, vol. 75, no. 3, pp. 289–299, 1982.

[4] A. Stoytchev, "Learning the affordances of tools using a behavior-grounded approach," in *Towards Affordance-Based Robot Control*, pp. 140–158, Springer, 2008.

[5] J. J. Gibson and R. Shaw, "Perceiving, acting, and knowing: Toward an ecological psychology," *The Theory of Affordances*, pp. 67–82, 1977.

[6] G. Rizzolatti, R. Camarda, L. Fogassi, M. Gentilucci, G. Luppino, and M. Matelli, "Functional organization of inferior area 6 in the macaque monkey," *Experimental brain research*, vol. 71, no. 3, pp. 491–507, 1988.

[7] P. Cisek, "Cortical mechanisms of action selection: the affordance competition hypothesis," *Philosophical Transactions of the Royal Society B: Biological Sciences*,

vol. 362, no. 1485, pp. 1585–1599, 2007.

[8] G. Cotugno, K. Althoefer, and T. Nanayakkara, "The role of the thumb: Study of finger motion in grasping and reachability space in human and robotic hands," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. PP, no. 99, pp. 1–10, 2016.

[9] G. Salvietti, M. Malvezzi, G. Gioioso, and D. Prattichizzo, "On the use of homogeneous transformations to map human hand movements onto robotic hands," in *Proc. IEEE Int. Conf. on Robotics and Automation, no. 0,(Hong Kong, China)*, 2014.

[10] R. Detry, E. Baseski, M. Popovic, Y. Touati, N. Kruger, O. Kroemer, J. Peters, and J. Piater, "Learning object-specific grasp affordance densities," in *2009 IEEE 8th International Conference on Development and Learning*, pp. 1–7, June 2009.

[11] W. J. Baumol, "Productivity growth, convergence, and welfare: what the long-run data show," *The American Economic Review*, pp. 1072–1085, 1986.

[12] M. A. C. C. of Afghanistan staff, "Mine clearance techniques," tech. rep., UN Afghan Mine Action Standards, March 2011.

[13] S. Sakakibara, B. B. Flynn, R. G. Schroeder, and W. T. Morris, "The impact of just-in-time manufacturing and its infrastructure on manufacturing performance," *Management Science*, vol. 43, no. 9, pp. 1246–1257, 1997.

[14] K. Marx, *Economic and Philosophic Manuscripts*. New York, International Publishers, 1844.

[15] M. W. Marzke and R. F. Marzke, "Evolution of the human hand: approaches to acquiring, analysing and interpreting the anatomical evidence.," *Journal of*

anatomy, vol. 197 (Pt 1), pp. 121–40, July 2000.

[16] T. Doege and T. Houston, *Guides to the evaluation of permanent impairment 3rd ed.* American Medical Association Press, 1993.

[17] M. Haslam, A. Hernandez-Aguilar, V. Ling, S. Carvalho, I. de La Torre, A. DeStefano, A. Du, B. Hardy, J. Harris, L. Marchant, *et al.*, "Primate archaeology," *Nature*, vol. 460, no. 7253, pp. 339–344, 2009.

[18] E. Pouydebat, M. Laurin, P. Gorce, and V. Bels, "Evolution of grasping among anthropoids.," *Journal of evolutionary biology*, vol. 21, pp. 1732–43, Nov. 2008.

[19] R. W. Young, "Evolution of the human hand: the role of throwing and clubbing," *Journal of Anatomy*, vol. 202, no. 1, pp. 165–174, 2003.

[20] B. R. Benefit and M. L. McCrossin, "Earliest known old world monkey skull," *Nature*, vol. 388, no. 6640, p. 368, 1997.

[21] R. Tomovic and G. Boni, "An adaptive artificial hand," *IRE Transactions on Automatic Control*, vol. 7, no. 3, pp. 3–10, 1962.

[22] M. Cutkosky, "On grasp choice, grasp models, and the design of hands for manufacturing tasks," *IEEE Transactions on Robotics and Automation*, vol. 5, pp. 269–279, June 1989.

[23] J. K. Salisbury and J. J. Craig, "Articulated hands: Force control and kinematic issues," *The International journal of Robotics research*, vol. 1, no. 1, pp. 4–17, 1982.

[24] G. Lundstrom, "Industrial robot grippers," *Industrial Robot: An International Journal*, vol. 1, no. 2, pp. 72–82, 1974.

[25] L. Jamone, E. Ugur, A. Cangelosi, L. Fadiga, A. Bernardino, J. Piater, and J. Santos-Victor, "Affordances in psychology, neuroscience and robotics: a survey," *IEEE Transactions on Cognitive and Developmental Systems*, vol. PP, no. 99, pp. 1–1, 2016.

[26] W. Spalteholz, *Atlas of human anatomy*. Butterworth-Heinemann, 2013.

[27] T. M. Greiner, "Hand anthropometry of U.S. army personnel," tech. rep., United States Army Natick Research, Development and Engineering Center, 1991.

[28] H. Gray, *Gray's anatomy: with original illustrations by Henry Carter*. Arcturus Publishing, 2009.

[29] R. Z. LeGeros, "Properties of osteoconductive biomaterials: calcium phosphates.," *Clinical orthopaedics and related research*, vol. 395, pp. 81–98, 2002.

[30] B. G. de Monsabert, J. Rossi, E. Berton, and L. Vigouroux, "Quantification of hand and forearm muscle forces during a maximal power grip task," *Medicine and science in sports and exercise*, vol. 44, no. 10, pp. 1906–1916, 2012.

[31] E. N. Marieb and K. Hoehn, *Human anatomy & physiology*. Pearson Education, 2007.

[32] D. H. ELLIOTT, "Structure and function of mammalian tendon," *Biological Reviews*, vol. 40, no. 3, pp. 392–421, 1965.

[33] R. E. Shadwick, "Elastic energy storage in tendons: mechanical differences related to function and age," *Journal of applied physiology*, vol. 68, no. 3, pp. 1033–1040, 1990.

[34] Y. Kawakami, T. Abe, H. Kanehisa, and T. Fukunaga, "Human skeletal mus-

cle size and architecture: variability and interdependence," *American Journal of Human Biology*, vol. 18, no. 6, pp. 845–848, 2006.

[35] K. Kuczynski, "Carpometacarpal joint of the human thumb.," *Journal of anatomy*, vol. 118, no. Pt 1, p. 119, 1974.

[36] W. P. Cooney, M. J. Lucca, E. Chao, and R. Linscheid, "The kinesiology of the thumb trapeziometacarpal joint.," *The Journal of Bone and Joint Surgery*, vol. 63, no. 9, pp. 1371–1381, 1981.

[37] E. T. Emerson, T. J. Krizek, and D. P. Greenwald, "Anatomy, physiology, and functional restoration of the thumb," *Annals of plastic surgery*, vol. 36, no. 2, pp. 180–191, 1996.

[38] T. Feix, R. Pawlik, H. Schmiedmayer, J. Romero, and D. Kragic, "A comprehensive grasp taxonomy," in *Robotics, Science and Systems: Workshop on Understanding the Human Hand for Advancing Robotic Manipulation*, June 2009.

[39] J. R. Napier, "The prehensile movements of the human hand," *Journal of bone and joint surgery*, vol. 38, no. 4, pp. 902–913, 1956.

[40] M. Brown and P. Matthews, "On the subdivision of the efferent fibres to muscle spindles into static and dynamic fusimotor fibres," *Control and innervation of skeletal muscle*, pp. 18–31, 1966.

[41] M. Hulliger, "The mammalian muscle spindle and its central control," in *Reviews of Physiology, Biochemistry and Pharmacology, Volume 101*, pp. 1–110, Springer, 1984.

[42] K. Cole, V. Gracco, and J. Abbs, "Autogenic and nonautogenic sensorimotor actions in the control of multiarticulate hand movements," *Experimental Brain*

Research, vol. 56, no. 3, pp. 582–585, 1984.

[43] F. Baldissera, H. Hultborn, and M. Illert, "Integration in spinal neuronal systems," *Comprehensive Physiology*, 1981.

[44] M. Santello, M. Flanders, and J. F. Soechting, "Postural hand synergies for tool use," *The Journal of Neuroscience*, vol. 18, no. 23, pp. 10105–10115, 1998.

[45] A. VALLBO, "Basic patterns of muscle spindle discharge," in *Muscle Receptors and Movement: Proceedings of a Symposium held at the Sherrington School of Physiology, St Thomas's Hospital Medical School, London, on July 8th and 9th, 1980*, p. 263, Springer, 1981.

[46] P. E. Crago, J. C. Houk, and W. Z. Rymer, "Sampling of total muscle force by tendon organs," *Journal of neurophysiology*, vol. 47, no. 6, pp. 1069–1083, 1982.

[47] R. Llinas, "Functional significance of the basic cerebellar circuit in motor coordination," in *Cerebellar functions*, pp. 170–185, Springer, 1984.

[48] E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, and A. J. Hudspeth, *Principles of neural science*, vol. 4. McGraw-hill New York, 2000.

[49] M. Ito, "Long-term depression as a memory process in the cerebellum," *Neuroscience Research*, vol. 3, no. 6, pp. 531 – 539, 1986.

[50] J. S. Albus, "A theory of cerebellar function," *Mathematical Biosciences*, vol. 10, no. 1-2, pp. 25–61, 1971.

[51] T. Bonhoeffer, V. Staiger, and A. Aertsen, "Synaptic plasticity in rat hippocampal slice cultures: local "hebbian" conjunction of pre-and postsynaptic stimulation leads to distributed synaptic enhancement," *Proceedings of the National Academy*

of Sciences, vol. 86, no. 20, pp. 8113–8117, 1989.

[52] C. Luscher, R. Nicoll, R. Malenka, and D. Muller, "Synaptic plasticity and dynamic modulation of the postsynaptic membrane," *Nature neuroscience*, vol. 3, no. 6, p. nn0600 545, 2000.

[53] L. E. White, *Medical Neuroscience Online Lectures*. Coursera, 2014.

[54] E. V. Evarts, "Relation of pyramidal tract activity to force exerted during voluntary movement," *J Neurophysiol*, vol. 31, no. 1, pp. 14–27, 1968.

[55] L. Deecke, P. Scheid, and H. H. Kornhuber, "Distribution of readiness potential, pre-motion positivity, and motor potential of the human cerebral cortex preceding voluntary finger movements," *Experimental Brain Research*, vol. 7, no. 2, pp. 158–168, 1969.

[56] M. Jeannerod, M. A. Arbib, G. Rizzolatti, and H. Sakata, "Grasping objects: the cortical mechanisms of visuomotor transformation," *Trends in neurosciences*, vol. 18, no. 7, pp. 314–320, 1995.

[57] O. Hikosaka, M. Matsumura, J. Kojima, and T. Gardiner, "Role of basal ganglia in initiation and suppression of saccadic eye movements," *Role of the cerebellum and basal ganglia in voluntary movement*, pp. 213–219, 1993.

[58] R. L. Albin, A. B. Young, and J. B. Penney, "The functional anatomy of basal ganglia disorders," *Trends in neurosciences*, vol. 12, no. 10, pp. 366–375, 1989.

[59] G. Rizzolatti, L. Fogassi, and V. Gallese, "Parietal cortex: from sight to action," *Current opinion in neurobiology*, vol. 7, no. 4, pp. 562–567, 1997.

[60] A. H. Fagg and M. A. Arbib, "Modeling parietal–premotor interactions in primate

control of grasping," *Neural Networks*, vol. 11, no. 7, pp. 1277–1303, 1998.

[61] E. P. Gardner, K. S. Babu, S. Ghosh, A. Sherwood, and J. Chen, "Neurophysiology of prehension iii. representation of object features in posterior parietal cortex of the macaque monkey," *Journal of neurophysiology*, vol. 98, no. 6, pp. 3708–3730, 2007.

[62] A. Schmitz, U. Pattacini, F. Nori, L. Natale, G. Metta, and G. Sandini, "Design, realization and sensorization of the dexterous iCub hand," in *10th IEEE-RAS International Conference on Humanoid Robots*, pp. 186–191, Ieee, Dec. 2010.

[63] H. Liu, P. Meusel, N. Seitz, B. Willberg, G. Hirzinger, M. Jin, Y. Liu, R. Wei, and Z. Xie, "The modular multisensory dlr-hit-hand," *Mechanism and Machine Theory*, vol. 42, no. 5, pp. 612 – 625, 2007.

[64] N. Dechev, W. Cleghorn, and S. Naumann, "Multiple finger, passive adaptive grasp prosthetic hand," *Mechanism and Machine Theory*, vol. 36, no. 10, pp. 1157 – 1173, 2001.

[65] T. LalibertÃ© and C. M. Gosselin, "Simulation and design of underactuated mechanical hands," *Mechanism and Machine Theory*, vol. 33, no. 1, pp. 39 – 57, 1998.

[66] C. Brown and H. Asada, "Inter-finger coordination and postural synergies in robot hands via mechanical implementation of principal components analysis," in *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*, pp. 2877–2882, 2007.

[67] M. G. Catalano, G. Grioli, E. Farnioli, A. Serio, C. Piazza, and A. Bicchi, "Adaptive synergies for the design and control of the pisa/iit softhand," *The International Journal of Robotics Research*, vol. 33, no. 5, pp. 768–782, 2014.

[68] S. Jacobsen, E. Iversen, D. Knutti, R. Johnson, and K. Biggers, "Design of the utah/mit dextrous hand," in *Robotics and Automation. Proceedings. 1986 IEEE International Conference on*, vol. 3, pp. 1520–1532, IEEE, 1986.

[69] W. Townsend, "The BarrettHand grasper - programmably flexible part handling and assembly," *Industrial Robot: An International Journal*, vol. 27, no. 3, pp. 181–188, 2000.

[70] "Shadow Dexterous Hand E1 Series, Technical Specification," Tech. Rep. January, Shadow Company, 2013.

[71] G. Wei, J. S. Dai, S. Wang, and H. Luo, "Kinematic analysis and prototype of a metamorphic anthropomorphic hand with a reconfigurable palm," *International Journal of Humanoid Robotics*, vol. 08, no. 03, pp. 459–479, 2011.

[72] S. Hirose and Y. Umetani, "The development of soft gripper for the versatile robot hand," *Mechanism and Machine Theory*, vol. 13, no. 3, pp. 351 – 359, 1978.

[73] M. Bianchi, P. Salaris, and a. Bicchi, "Synergy-based hand pose sensing: Optimal glove design," *The International Journal of Robotics Research*, vol. 32, pp. 407–424, Apr. 2013.

[74] E. P. Pitarch, *Virtual Human Hand : Grasping Strategy and Simulation.* PhD thesis, Universitat Politecnica de Catalunya, 2007.

[75] V. K. Nanayakkara, G. Cotugno, N. Vitzilaios, D. Venetsanos, T. Nanayakkara, and M. N. Sahinkaya, "The role of morphology of the thumb in anthropomorphic grasping: a review," *Frontiers in Mechanical Engineering*, vol. 3, no. 5, 2017.

[76] M. Grebenstein, *Approaching Human Performance*, ch. Analysis of the Human Hand. 98, Springer International Publishing, 2014.

[77] C. Melchiorri, G. Palli, G. Berselli, and G. Vassura, "Development of the ub hand iv: Overview of design solutions and enabling technologies," *IEEE Robotics Automation Magazine*, vol. 20, pp. 72–81, Sept 2013.

[78] M. Malosio, G. Spagnuolo, A. Prini, L. M. Tosatti, and G. Legnani, "Principle of operation of rotwwc-vsa, a multi-turn rotational variable stiffness actuator," *Mechanism and Machine Theory*, vol. 116, pp. 34–49, 2017.

[79] J. Fras, M. Macias, F. Czubaczynski, P. Salek, and J. Glowka, "Soft flexible gripper design, characterization and application," in *International Conference on Systems, Control and Information Technologies 2016*, pp. 368–377, Springer, 2016.

[80] R. Deimel and O. Brock, "A novel type of compliant and underactuated robotic hand for dexterous grasping," *The International Journal of Robotics Research*, vol. 35, no. 1-3, pp. 161–185, 2016.

[81] D. Rus and M. T. Tolley, "Design, fabrication and control of soft robots," *Nature*, vol. 521, no. 7553, pp. 467–475, 2015.

[82] M. Grebenstein, M. Chalon, G. Hirzinger, and R. Siegwart, "Antagonistically driven finger design for the anthropomorphic dlr hand arm system," in *2010 10th IEEE-RAS International Conference on Humanoid Robots*, pp. 609–616, Dec 2010.

[83] H. Kawasaki, T. Komatsu, and K. Uchiyama, "Dexterous anthropomorphic robot hand with distributed tactile sensor: Gifu hand II," *IEEE/ASME Transactions on Mechatronics*, vol. 7, pp. 296–303, Sept. 2002.

[84] G. Cotugno, K. Konstantinova, A. Stilli, Y. Noh, and K. Althoefer, "Object classification using hybrid fiber optical force/proximity sensor," in *IEEE Sensors Conference*, 2017.

[85] W. Penfield and T. Rasmussen, *The cerebral cortex of man: a clinical study of localization of function.* Oxford, England: Macmillan, 1950.

[86] A. Leo, G. Bernardi, G. Handjaras, D. Bonino, E. Ricciardi, and P. Pietrini, "Increased bold variability in the parietal cortex and enhanced parieto-occipital connectivity during tactile perception in congenitally blind individuals," *Neural plasticity*, vol. 2012, 2012.

[87] H. Mellmann and G. Cotugno, "Dynamic motion control: Adaptive bimanual grasping for a humanoid robot," *Fundamenta Informaticae*, vol. 112, no. 1, pp. 89–101, 2011.

[88] A. Del Prete, S. Denei, L. Natale, F. Mastrogiovanni, F. Nori, G. Cannata, and G. Metta, "Skin Spatial Calibration Using Force / Torque Measurements," in *Intelligent Robots and Systems (IROS)*, 2011.

[89] M. Quigley, C. Salisbury, A. Y. Ng, and J. K. Salisbury, "Mechatronic design of an integrated robotic hand," *The International Journal of Robotics Research*, vol. 33, no. 5, pp. 706–720, 2014.

[90] J. Konstantinova, A. Stilli, and K. Althoefer, "Force and proximity fingertip sensor to enhance grasping perception," in *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*, pp. 2118–2123, IEEE, 2015.

[91] J. Konstantinova, S. Krivic, A. Stilli, J. Piater, and K. Althoefer, "Autonomous object handover using wrist tactile information," in *Conference Towards Autonomous Robotic Systems*, pp. 450–463, Springer, 2017.

[92] D. Novak and R. Riener, "A survey of sensor fusion methods in wearable robotics," *Robotics and Autonomous Systems*, vol. 73, no. Supplement C, pp. 155 – 170, 2015.

[93] M. T. Mason and J. K. Salisbury, *Robot Hands and the Mechanics of Manipulation.* The MIT series in Artificial Intelligence, Cambridge, Massachusetts: The MIT Press, 1985.

[94] A. Bicchi, "On the closure properties of robotic grasping," *The International Journal of Robotics Research*, vol. 14, no. 4, pp. 319–334, 1995.

[95] K. B. Shimoga, "Robot grasp synthesis algorithms: A survey," *The International Journal of Robotics Research*, vol. 15, no. 3, pp. 230–266, 1996.

[96] M. Kopicki, R. Detry, F. Schmidt, C. Borst, R. Stolkin, and J. L. Wyatt, "Learning dexterous grasps that generalise to novel objects by combining hand and contact models," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5358–5365, IEEE, 2014.

[97] A. Bicchi and V. Kumar, "Robotic grasping and contact: A review," in *Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference on*, vol. 1, pp. 348–353, IEEE, 2000.

[98] A. Bicchi, M. Gabiccini, and M. Santello, "Modelling natural and artificial hands with synergies," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 366, no. 1581, pp. 3153–3161, 2011.

[99] G. Grioli, M. Catalano, E. Silvestro, S. Tono, and A. Bicchi, "Adaptive synergies: an approach to the design of under-actuated robotic hands," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pp. 1251–1256, IEEE, 2012.

[100] M. Pozzi, A. M. Sundaram, M. Malvezzi, D. Prattichizzo, and M. A. Roa, "Grasp quality evaluation in underactuated robotic hands," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1946–1953, Oct

2016.

[101] D. Prattichizzo, M. Malvezzi, M. Gabiccini, and A. Bicchi, "On motion and force controllability of precision grasps with hands actuated by soft synergies," *IEEE Transactions on Robotics*, vol. 29, pp. 1440–1456, Dec 2013.

[102] S. Micera, L. Citi, J. Rigosa, J. Carpaneto, S. Raspopovic, G. D. Pino, L. Rossini, K. Yoshida, L. Denaro, P. Dario, and P. M. Rossini, "Decoding information from neural signals recorded using intraneural electrodes: Toward the development of a neurocontrolled hand prosthesis," *Proceedings of the IEEE*, vol. 98, pp. 407–417, March 2010.

[103] G. Cotugno, V. Mohan, K. Althoefer, and T. Nanayakkara, "Simplifying grasping complexity through generalization of kinaesthetically learned synergies," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5345–5351, IEEE, 2014.

[104] H. Ben Amor, O. Kroemer, U. Hillenbrand, G. Neumann, and J. Peters, "Generalization of human grasping for multi-fingered robot hands," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pp. 2043–2050, 2012.

[105] T. Hsia, "Adaptive control of robot manipulators - a review," in *Proceedings. 1986 IEEE International Conference on Robotics and Automation*, vol. 3, pp. 183–189, Apr 1986.

[106] D. Katz, *Interactive perception of articulated objects for autonomous manipulation.* PhD thesis, University of Massachusetts Amherst, 2011.

[107] F. Zhang, J. Leitner, M. Milford, B. Upcroft, and P. Corke, "Towards vision-based deep reinforcement learning for robotic motion control," *arXiv preprint*

arXiv:1511.03791, 2015.

[108] S. Gu, E. Holly, T. Lillicrap, and S. Levine, "Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3389–3396, May 2017.

[109] C. Y. Brown and H. H. Asada, "Inter-finger coordination and postural synergies in robot hands via mechanical implementation of principal components analysis," *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2877–2882, Oct. 2007.

[110] S. Schaal, J. Peters, J. Nakanishi, and A. Ijspeert, "Learning movement primitives," *Robotics Research*, pp. 561–572, 2005.

[111] O. Kroemer, E. Ugur, E. Oztop, and J. Peters, "A kernel-based approach to direct action perception.," in *ICRA*, pp. 2605–2610, IEEE, 2012.

[112] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Robotics and Autonomous Systems*, 2009.

[113] C. L. Nehaniv and K. Dautenhahn, "2 the correspondence problem," *Imitation in animals and artifacts*, p. 41, 2002.

[114] P. Kremer, T. Wimbock, J. Artigas, S. Schatzle, K. Johl, F. Schmidt, C. Preusche, and G. Hirzinger, "Multimodal telepresent control of dlr's rollin justin," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*, pp. 1601–1602, IEEE, 2009.

[115] T. Feix, J. Romero, C. H. Ek, H.-B. Schmiedmayer, and D. Kragic, "A Metric for Comparing the Anthropomorphic Motion Capability of Artificial Hands," *IEEE*

Transactions on Robotics, vol. 29, pp. 82–93, Feb. 2013.

[116] J. Peters, D. D. Lee, J. Kober, D. Nguyen-Tuong, J. A. Bagnell, and S. Schaal, "Robot learning," in *Springer Handbook of Robotics*, pp. 357–398, Springer, 2016.

[117] J. Konstantinova, M. Li, V. Aminzadeh, P. Dasgupta, K. Althoefer, and T. Nanayakkara, "Force-velocity modulation strategies for soft tissue examination," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1998–2003, Nov 2013.

[118] S. Puhlmann, F. Heinemann, O. Brock, and M. Maertens, "A compact representation of human single-object grasping," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1954–1959, Oct 2016.

[119] Y. Yang and O. Brock, "Elastic roadmaps-motion generation for autonomous mobile manipulation," *Autonomous Robots*, vol. 28, no. 1, pp. 113–130, 2010.

[120] O. Khatib, "Real-time obstacle avoidance for manipulators and mobile robots," *The international journal of robotics research*, vol. 5, no. 1, pp. 90–98, 1986.

[121] V. Mohan and P. Morasso, "Passive motion paradigm: an alternative to optimal control," *Frontiers in Neurorobotics*, vol. 5, no. 4, 2011.

[122] M. Leonetti, L. Iocchi, and P. Stone, "A synthesis of automated planning and reinforcement learning for efficient, robust decision-making," *Artificial Intelligence*, vol. 241, pp. 103–130, 2016.

[123] E. Sahin, M. Cakmak, M. R. Dogar, E. Ugur, and G. Ucoluk, "To afford or not to afford: A new formalization of affordances toward affordance-based robot control," *Adaptive Behavior*, vol. 15, no. 4, pp. 447–472, 2007.

[124] L. Montesano, M. Lopes, A. Bernardino, and J. Santos-Victor, "Learning object affordances: From sensory–motor coordination to imitation," *IEEE Transactions on Robotics*, vol. 24, no. 1, pp. 15–26, 2008.

[125] V. Tikhanoff, U. Pattacini, L. Natale, and G. Metta, "Exploring affordances and tool use on the icub," in *2013 13th IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, pp. 130–137, Oct 2013.

[126] A. Goncalves, J. Abrantes, G. Saponaro, L. Jamone, and A. Bernardino, "Learning intermediate object affordances: Towards the development of a tool concept," in *Development and Learning and Epigenetic Robotics (ICDL-Epirob), 2014 Joint IEEE International Conferences on*, pp. 482–488, IEEE, 2014.

[127] P. Fitzpatrick, G. Metta, L. Natale, S. Rao, and G. Sandini, "Learning about objects through action - initial steps towards artificial cognition," in *2003 IEEE International Conference on Robotics and Automation (Cat. No.03CH37422)*, vol. 3, pp. 3140–3145 vol.3, Sept 2003.

[128] T. Hermans, F. Li, J. M. Rehg, and A. F. Bobick, "Learning contact locations for pushing and orienting unknown objects," in *2013 13th IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, pp. 435–442, Oct 2013.

[129] A. Galuszka and A. Swierniak, "Planning in multi-agent environment using strips representation and non-cooperative equilibrium strategy," *Journal of Intelligent and Robotic Systems*, vol. 58, no. 3, pp. 239–251, 2010.

[130] V. A. Ziparo, L. Iocchi, P. Lima, D. Nardi, and P. Palamara, "Petri net plans," in *Proceedings of Fourth International Workshop on Modelling of Objects, Components, and Agents (MOCA)*, pp. 267–290, 2006.

[131] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare,

A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[132] M. V. Weghe, M. Rogers, M. Weissert, and Y. Matsuoka, "The act hand: Design of the skeletal structure," in *Robotics and Automation, 2004. Proceedings. ICRA2004. 2004 IEEE International Conference on*, vol. 4, pp. 3375–3379, IEEE, 2004.

[133] H. Reimann, I. Iossifidis, and G. Schoner, "Generating collision free reaching movements for redundant manipulators using dynamical systems," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pp. 5372–5379, Oct 2010.

[134] S. Kim, A. Shukla, and A. Billard, "Catching objects in flight," *IEEE Transactions on Robotics*, vol. 30, pp. 1049–1065, Oct 2014.

[135] U. Castiello, K. M. Bennett, and Y. Paulignan, "Does the type of prehension influence the kinematics of reaching," *Behavioural brain research*, vol. 50, no. 1, pp. 7–15, 1992.

[136] M. Gentilucci, E. Daprati, M. Gangitano, M. C. Saetti, and I. Toni, "On orienting the hand to reach and grasp an object.," *Neuroreport*, vol. 7, no. 2, pp. 589–592, 1996.

[137] K. Hausman, S. Niekum, S. Osentoski, and G. S. Sukhatme, "Active articulation model estimation through interactive perception," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3305–3312, May 2015.

[138] Y. Caliskan, P. Duygulu, and E. Sahin, "Affordance prediction of hand tools using interactive perception," in *2012 20th Signal Processing and Communications Applications Conference (SIU)*, pp. 1–4, April 2012.

[139] L. Sun, S. Rogers, G. Aragon-Camarasa, and J. P. Siebert, "Recognising the clothing categories from free-configuration using gaussian-process-based interactive perception," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2464–2470, May 2016.

[140] J. Aloimonos, I. Weiss, and A. Bandyopadhyay, "Active vision," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 333–356, 1988.

[141] E. Arruda, J. Wyatt, and M. Kopicki, "Active vision for dexterous grasping of novel objects," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2881–2888, Oct 2016.

[142] B. Calli, M. Wisse, and P. Jonker, "Grasping of unknown objects via curvature maximization using active vision," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 995–1001, Sept 2011.

[143] L. Fenson, J. Kagan, R. B. Kearsley, and P. R. Zelazo, "The developmental progression of manipulative play in the first two years," *Child Development*, vol. 47, no. 1, pp. 232–236, 1976.

[144] N. Lawrence, "Mocap toolbox for matlab," *Available on-line at http://www. cs. man. ac. uk/neill/mocap*, 2011.

[145] J. Lin, Y. Wu, and T. S. Huang, "Modeling the constraints of human hand motion," in *Human Motion, 2000. Proceedings. Workshop on*, pp. 121–126, IEEE, 2000.

[146] T. Flash and N. Hogan, "The coordination of arm movements: an experimentally confirmed mathematical model," *The journal of Neuroscience*, vol. 5, no. 7, pp. 1688–1703, 1985.

[147] V. Gaveau, L. Pisella, A.-E. Priot, T. Fukui, Y. Rossetti, D. Pelisson, and

C. Prablanc, "Automatic online control of motor adjustments in reaching and grasping," *Neuropsychologia*, vol. 55, pp. 25–40, 2014.

[148] J. B. J. Smeets and E. Brenner, *Synergies in Grasping*, pp. 21–34. Cham: Springer International Publishing, 2016.

[149] D. Bullock and S. Grossberg, "Neural dynamics of planned arm movements: emergent invariants and speed-accuracy properties during trajectory formation.," *Psychological review*, vol. 95, no. 1, p. 49, 1988.

[150] R. Marteniuk, C. MacKenzie, M. Jeannerod, S. Athenes, and C. Dugas, "Constraints on human arm movement trajectories.," *Canadian Journal of Psychology/Revue canadienne de psychologie*, vol. 41, no. 3, p. 365, 1987.

[151] M. Jeannerod, "The Timing of Natural Prehension Movements," *Journal of Motor Behavior*, vol. 16, pp. 235–254, Sept. 1984.

[152] H. S. Milner-Brown, R. B. Stein, and R. Yemm, "The contractile properties of human motor units during voluntary isometric contractions," *The Journal of Physiology*, vol. 228, no. 2, pp. 285–306, 1973.

[153] M. W. Marzke, "Precision grips, hand morphology, and tools," *American Journal of Physical Anthropology*, vol. 102, no. 1, pp. 91–110, 1997.

[154] I. Kapandji, *The Physiology of the Joints: Upper limb*. Physiology of the Joints, Churchill Livingstone, 1982.

[155] EU FP7 GRASP Project Consortium, "Human Grasping Database." `http://grasp.xief.net/`, September 2014.

[156] M. Diftler, J. Mehling, M. Abdallah, N. Radford, L. Bridgwater, A. Sanders,

R. Askew, D. Linn, J. Yamokoski, F. Permenter, B. Hargrave, R. Piatt, R. Savely, and R. Ambrose, "Robonaut 2 - the first humanoid robot in space," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pp. 2178–2183, May 2011.

[157] G. Wei, J. S. Dai, S. Wang, and H. Luo, "Kinematic analysis and prototype of a metamorphic anthropomorphic hand with a reconfigurable palm," *International Journal of Humanoid Robotics*, vol. 8, no. 03, pp. 459–479, 2011.

[158] J. Denavit and R. S. Hartenberg, "A kinematic notation for lower-pair mechanisms based on matrices," *Journal of Applied Mechanics*, 1955.

[159] L. Cui, U. Cupcic, and J. S. Dai, "An optimization approach to teleoperation of the thumb of a humanoid robot hand: Kinematic mapping and calibration," *Journal of Mechanical Design*, vol. 136, no. 9, p. 091005, 2014.

[160] RobotCub Wiki, "iCub Right Hand Forward Kinematics." `http://wiki.icub.org/wiki/ICubFowardKinematics_hand_right`, September 2014.

[161] I. E. Sutherland and G. W. Hodgman, "Reentrant polygon clipping," *Communications of the ACM*, vol. 17, no. 1, pp. 32–42, 1974.

[162] U. Prieur, V. Perdereau, and A. Bernardino, "Modeling and planning high-level in-hand manipulation actions from human knowledge and active learning from demonstration," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pp. 1330–1336, IEEE, 2012.

[163] A. J. Ijspeert, J. Nakanishi, and S. Schaal, "Movement imitation with nonlinear dynamical systems in humanoid robots," in *Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on*, vol. 2, pp. 1398–1403, IEEE, 2002.

[164] D. Consortium, "Deliverable d5.1: Body models of the relevant robotic platforms," tech. rep., Italian Institute of Technology, July 2012.

[165] R. L. Williams II, "Baxter humanoid robot kinematics," tech. rep., Dept. Mechanical Engineering, Ohio University, April 2017.

[166] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, "Ros: an open-source robot operating system," in *ICRA workshop on open source software*, vol. 3, p. 5, Kobe, 2009.

[167] W. Kahan, "Pracniques: further remarks on reducing truncation errors," *Communications of the ACM*, vol. 8, no. 1, p. 40, 1965.

[168] S. M. LaValle and J. J. Kuffner Jr, "Rapidly-exploring random trees: Progress and prospects," *Citeseer*, 2000.

[169] J. J. Kuffner and S. M. LaValle, "Rrt-connect: An efficient approach to single-query path planning," in *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No.00CH37065)*, vol. 2, pp. 995–1001 vol.2, 2000.

[170] I. Sucan, M. Moll, and E. Kavraki, "The open motion planning library," *Robotics and Automation Magazine, IEEE*, vol. 19, pp. 72–82, 12 2012.

[171] I. A. Sucan and S. Chitta, "Moveit!," *Online at http://moveit. ros. org*, 2013.

[172] C. Mason, J. Gomez, and T. Ebner, "Hand synergies during reach-to-grasp," *Journal of Neurophysiology*, vol. 86, no. 6, pp. 2896–2910, 2001.

[173] V. Klema and A. Laub, "The singular value decomposition: Its computation and some applications," *Automatic Control, IEEE Transactions on*, vol. 25, no. 2,

pp. 164–176, 1980.

[174] K. Althoefer, B. Krekelberg, D. Husmeier, and L. Seneviratne, "Reinforcement learning in a rule-based navigator for robotic manipulators," *Neurocomputing*, vol. 37, no. 1-4, pp. 51–70, 2001.

[175] H. Ben Amor, O. Kroemer, U. Hillenbrand, G. Neumann, and J. Peters, "Generalization of human grasping for multi-fingered robot hands," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pp. 2043–2050, IEEE, 2012.

[176] A. Savitzky and M. J. E. Golay, "Smoothing and Differentiation of Data by Simplified Least Squares Procedures," *Analytical Chemistry*, vol. 36, 1964.

[177] M. Santello, M. Flanders, and J. F. Soechting, "Patterns of hand motion during grasping and the influence of sensory guidance," *The Journal of Neuroscience*, vol. 22, no. 4, pp. 1426–1435, 2002.

[178] J. B. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. of the fifth Berkeley Symposium on Mathematical Statistics and Probability* (L. M. L. Cam and J. Neyman, eds.), vol. 1, pp. 281–297, University of California Press, 1967.

[179] E. Chinellato, R. Fisher, A. Morales, and A. Del Pobil, "Ranking planar grasp configurations for a three-finger hand," in *Robotics and Automation, 2003. Proceedings. ICRA '03. IEEE International Conference on*, vol. 1, pp. 1133–1138 vol.1, 2003.

[180] G. Metta, L. Natale, F. Nori, G. Sandini, D. Vernon, L. Fadiga, C. von Hofsten, K. Rosander, M. Lopes, J. Santos-Victor, A. Bernardino, and L. Montesano, "The icub humanoid robot: An open-systems platform for research in cognitive

development," *Neural Networks*, vol. 23, no. 8-9, pp. 1125–1134, 2010.

[181] G. Cannata, M. Maggiali, G. Metta, and G. Sandini, "An embedded artificial skin for humanoid robots," in *Multisensor Fusion and Integration for Intelligent Systems, 2008. MFI 2008. IEEE International Conference on*, pp. 434 –438, 2008.

[182] P. Fitzpatrick, G. Metta, and L. Natale, "Towards long-lived robot genes," *Robot. Auton. Syst.*, vol. 56, no. 1, pp. 29–45, 2008.

[183] B. Siciliano and O. Khatib, *Springer handbook of robotics*. Springer, 2008.

[184] M. W. Marzke and K. L. Wullstein, "Chimpanzee and human grips: A new classification with a focus on evolutionary morphology," *International Journal of Primatology*, vol. 17, no. 1, pp. 117–139, 1996.

[185] M. A. Cusumano, "The limits of" lean"," *Sloan management review*, vol. 35, no. 4, p. 27, 1994.

[186] G. Pratt and J. Manzo, "The darpa robotics challenge [competitions]," *IEEE Robotics & Automation Magazine*, vol. 20, no. 2, pp. 10–12, 2013.

[187] B.-H. Juang and L. R. Rabiner, "Automatic speech recognition–a brief history of the technology development," *Georgia Institute of Technology. Atlanta Rutgers University and the University of California. Santa Barbara*, vol. 1, p. 67, 2005.

[188] R. R. Fay and L. A. Wilber, "Hearing in vertebrates: a psychophysics databook," *The Journal of the Acoustical Society of America*, vol. 86, no. 5, pp. 2044–2044, 1989.

[189] J. Neitz, T. Geist, and G. H. Jacobs, "Color vision in the dog," *Visual neuroscience*, vol. 3, no. 2, pp. 119–125, 1989.

[190] G. H. Jacobs, J. F. Deegan, M. A. Crognale, and J. A. Fenwick, "Photopigments of dogs and foxes and their implications for canid vision," *Visual neuroscience*, vol. 10, no. 1, pp. 173–180, 1993.

[191] Parmenides, "On nature," *De coelo*, V century B.C.

[192] M. Sprung, *The Question of Being: East-West Perspectives.* Pennsylvania State University Press, 1990.

[193] R. Descartes, *Discourse on the Method.* Leyde, de l'Imprimerie de Ian Maire, 1637.

[194] C. Li, *Toward a Contextual Approach to the Question of Being.* PhD thesis, The University of Connecticut, 1992.

[195] Socrates, *Teeto, Dialogues.* Plato, 365 BC.

Appendix A

Grasp Classification Tabular Summary

Posture class	GRASP name	Picture	Grasped Object
Power Grasp	Large diameter		Cylinder (11cm diameter)
Power Grasp	Small diameter		Cylinder (3cm diameter)
Power Grasp	Medium Wrap		Cylinder (3cm diameter)
Power Grasp	Power Disk		MiniDisc (8cm diameter, 2mm tall)
Power Grasp	Power Sphere		Tennis Ball (6.7cm diameter)
Power Grasp	Parallel Extension		Box (4cm thick)

Power Grasp	Lateral Tripod		Bottle Cap
Power Grasp	Sphere 4 Fingers		Tennis Ball (6.7cm diameter)
Power Grasp	Ring		Cylinder (6.4cm diameter)
Power Grasp	Inferior Pincer		Golf Ball (4.3cm diameter)
Precision Grasp	Prismatic 4 Fingers		Cylinder (1cm diameter)
Precision Grasp	Prismatic 3 Fingers		Cylinder (1cm diameter)
Precision Grasp	Prismatic 2 Fingers		Cylinder (1cm diameter)
Precision Grasp	Palmar Pinch		Coin
Precision Grasp	Precision Disk		Compact Disc
Precision Grasp	Precision Sphere		Tennis Ball (6.7cm diameter)
Precision Grasp	Tripod		Golf Ball (4.3cm diameter)
Precision Grasp	Tip Pinch		Cube 5mm
Precision Grasp	Quadpod		Golf Ball (4.3cm diameter)
Key Grasp	Adducted Thumb		Cylinder (3cm diameter)
Key Grasp	Light Tool		Cylinder (1cm diameter)

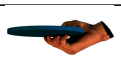
Key Grasp	Lateral		Credit Card
Key Grasp	Index Finger Extension		Cylinder (3cm diameter)
Key Grasp	Writing Tripod		Cylinder (1cm diameter)
Key Grasp	Sphere 3 Fingers		Tennis Ball (6.7cm diameter)
Key Grasp	Stick		Cylinder (1cm diameter)
Key Grasp	Ventral		Cylinder (1cm diameter)
Primate Grasp	Fixed Hook		Cylinder (3cm diameter)
Primate Grasp	Adduction Grip		Cylinder (1cm diameter)
Primate Grasp	Palmar		Plate
Not used	Extension Type		Plate

Table A.1: The table shows the mapping between the posture class defined in section 4.2.3 and the GRASP nomenclature as in the database [115]. For each entry of the GRASP database it is showed the GRASP nomenclature, the corresponding classification of this chapter, a pictorial representation of the posture (from the description of the GRASP online database [38]), the object grasped to produce the posture and the GRASP identification number used in the database.